Fragment based inhibitor design of *Mycobacterium tuberculosis* BioA

A Thesis
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

Ran Dai

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Barry. C. Finzel, Advisor

2015. 01

© Ran Dai 2015

Acknowledgements

First I would like to convey my sincere gratitude to my advisor, Dr. Barry. C. Finzel. He provided me with a lot of opportunities to explore in the field of drug discovery, especially to work on the BioA project presented in this dissertation. I would like to thank him for the guidance both for my academic growth and for my life. I am very glad that in the journey of obtaining the degree of Doctor of Philosophy, Barry taught me not only about science, but more importantly, about the philosophy of life. I made very clear goals for my life; and this is the most valuable thing I gained from my PhD study. I would also like to thank my committee members, Dr. Elizabeth Amin, Dr. Patrick Hanna, Dr. Courtney Aldrich and Dr. Hideki Aihara for their insightful comments and suggestions that contribute heavily to this dissertation.

The BioA would not have been possible without the help and support from the project team. I would like to specially thank Dr. Courtney Aldrich for initiating the collaboration, providing great discussions and suggestions. I would like to thank Dr. Todd Geders for his great contribution to this project, and for his very nice guidance to help me with techniques in molecular biology and crystallography. Thanks my collaborators Dr. Feng Liu and Dr. Daniel Wilson for the synthesis, ITC experiments and biochemical evaluations of the compounds.

The Finzel group has always been a very supportive family. I would like to thank the former members Dr. Todd geders, Dr. Teresa De la Mora for helping me getting started in my research; Dr. Jeffery Van Voorst and Dr. Kenneth Johnson for great suggestions and discussions. Thanks to my fellow students Li-Kai Liu and Kimberly Maize for both great discussions about research and being great companies in the lab. We have shared our excitements and they often cheer me up when I am down. Kim's bakery grade dessert is one of the very memorable thing in my graduate life.

Finally, I would like to thank my family and friends. I would like to thank my husband Cheng Zheng for being a supportive spouse as we went through this five-year long distance relationship. I would like to thank my mother Ling Li and my father Zhijun Dai for the support during the five years. Thank my mother Ling Li and mother-in-law Yuan Cai for helping me taking care of the baby while I am finishing my dissertation. I would like to thank my friends in University of Minnesota that made my life colorful here.

Dedication

This thesis is dedicated to my dear daughter Xina (Athena) Zheng, my husband Cheng Zheng, my mother Ling Li and my father Zhijun Dai

Abstract

7,8-Diaminopelargonic acid synthase (BioA) of *Mycobacterium tuberculosis* (*Mtb*) is a recently validated target for therapeutic intervention in the treatment of tuberculosis (TB). We herein report our fragment based inhibitor design of *Mtb* BioA. Using differential scanning fluorimetry (DSF) fragment screening, the Maybridge Ro3 library of 1000 molecules was screened. Twenty-one compounds giving rise to T_m shifts exceeding ±2°C were then investigated in crystallographic experiments. Six fragments have been co-crystallized with BioA to characterize binding. Each compound has a unique binding mode, and subtle variations in ligand binding site geometry are induced upon binding of different fragment molecules. Binding affinities of the fragments were characterized via isothermal titration calorimetry (ITC). A fragment extension strategy was used to rationally optimize these fragment hits. A commerce based SAR was used and identified 50 compounds containing the core of one of the fragments. These compounds were further screened virtually and experimently by DSF. Four optimized BioA ligands from fragment optimization were validated by X-ray crystallography, including a potent aryl hydrazine inhibitor of BioA that reversibly modifies the pyridoxal-5′-phosphate (PLP) cofactor. Binding affinities of these ligands have been characterized by ITC or kinetic assay. The six fragment complex structures were also used for optimization of HTS lead compounds. Six HTS lead compounds were co-crystallized with BioA at high resolution. Design of optimized compounds was by overlapping the fragments and HTS lead binding conformations in the BioA active site. Molecules predicted to have better potency were proposed. Two *N*-aryl piperazine inhibitors of BioA from HTS optimization were characterized using X-ray crystallography and ITC. One inhibitor that combines features of one HTS lead and one fragment was confirmed with improved binding affinity by ITC.

Table of Contents

List of Tables

List of Figures

Chapter I: Introduction

Acknowledgement: Figure 1.3 in this chapter is excerpted from previously published work [Dai, R. et al, 2014. "Inhibition of Mycobacterium tuberculosis Transaminase BioA by Aryl Hydrazines and Hydrazides ". ChemBioChem. 15(4):575-586].

1.1. Significance of novel anti-tubercular drugs

1.1.1. Current Progress on TB treatment

Tuberculosis (TB) is a major global health issue that causes over 1.8 million deaths annually[1]. *Mycobacterium tuberculosis* (*Mtb*) is the major pathogen causing TB in humans. *Mtb* first infects macrophages in the lungs. If the immune system fails to eradicate the bacilli in macrophages, *Mtb* can further spread to secondary infection sites in an uncontrolled pattern and lead to active clinical disease[2]. The *Mtb* infection can also be latent, in which case adaptive immunity is able to restrain the bacilli growth by controlling its replication for years, even decades, until a waning of immunity allows reactivation of the disease[3]. Based on the global incidence of a positive response in the tuberculin skin test, about 2 billion people have latent infection of *Mtb*. 10% of these 2 billion people will develop an active infection in their lifetime[1]. Because a compromised immune system can struggle to keep latent infections in check, HIV infection significantly increases the risk of conversion to active TB, and TB has become the major cause of death in people with AIDS[4].

The discovery of effective chemotherapies for TB began in the 1940s, and dozens of first and second line drugs have been developed since then. Let us review several landmarks in the history of TB medical development.

The first chemotherapies Streptomycin (STR) and *para*-aminosalicyclic acid (PAS) were discovered and used in the 1940s[5] (Figure 1.1a). STR belongs to the aminoglycoside antibiotic family. It is a protein synthesis inhibitor, which binds to the 30S subunit of the bacterial ribosome and causes codon misreading in bacterial protein synthesis. PAS was used as anti-mycobacterial for more then 60 years even though its mechanism of action remains only partially known. Recent studies have shown that PAS is a prodrug targeting dihydrofolate reductase (DHFR); it needs to be activated via the *Mtb* folate biosynthesis pathway to acquire DHFR inhibition activity[6,7].

In the 1950s, isoniazid (INH) was discovered and the triple drug combination (STR, PAS, INH) was found to be superior to other therapies (Figure 1.1 b1). Isoniazid is a prodrug activated by *Mtb* catalase-peroxidase (KatG) to form an isonicotinic acyl-NADH complex. This complex can bind to the enoyl-acyl carrier protein reductase (InhA) and block the binding of natural substrates. In this way, isoniazid inhibits the synthesis of mycolic acid, which is a crucial precursor for *Mtb* cell wall synthesis. In 1952, a triple therapy (streptomycin, para-aminosalicylic acid and isoniazid) resulting in an “assured cure” was established with a treatment duration of 18 months[5].

In the 1970s, the discovery of rifampicin (RIF) (Figure 1.1 b2) brought about a revolution in TB therapy, because it allowed the length of therapy to be significantly shortened (from 18 months to 9 months). Later in the 1980s, a 6-month short course chemotherapy was introduced for drug-susceptible TB that resulted in a cure rate exceeding 95%[8]. In this therapy, a four-drug regimen composed of RIF, INH, Ethambutol (EMB) (Figure 1.1 b3) and pyrazinamide (PZA) (Figure 1.1 b4) was used in the first 2 months as the intensive phase of therapy; then RIF and INH was used for an additional 4 months as a

continuation[9]. RIF, INH, EMB and PZA were subsequently named as first line anti-TB drugs. RIF is a DNA-dependent RNA polymerase inhibitor that inhibits *Mtb* RNA synthesis[9]. EMB disrupts *Mtb* cell wall synthesis by inhibiting an enzyme arabinosyl transferase, which is responsible in cell wall formation[10]. PZA is a prodrug; it first diffuses into *Mtb* and then it can be converted into the active metabolite pyrazinoic acid by pyrazinamidase; then the active form can inhibit *Mtb* fatty acid synthesis[11].

Beginning in the 1990s, strains of *Mtb* that are increasingly resistant to available drugs began to arise as formidable new challenges to human health. The term MultiDrug-Resistant TB (MDR-TB) is specifically used to refer to clinical isolates that are resistant to both INH and RIF, the two most important first line anti-TB drugs. To treat MDR-TB or TB with even higher degrees of resistance, WHO has recommended second line anti-TB drugs (WHO groups 2, 3, and 4) [8]. When treatment employing combinations of the four first line anti-TB drugs encounter resistance or intolerance issues, second line antibiotics should be used. These second line drugs include amikacin, kanamycin, capreomycin, fluoroquinolones, ethionamide, prothionamide, cycloserine, and PAS. These antibiotics are categorized as the second line drugs because they possess unconfirmed clinical efficacy, have substantially greater adverse effects, and in some cases unfavorable pharmacokinetic profiles. When fewer than 5 effective drugs from the first line and second line can be selected for the resistant TB, some third line antibiotics from a non-WHO approved list (macrolides such as clarithromycin, augmentin, meropenem and clavulanate, clofazimine, linezolid, thioacetazone, and thioridazine) are allowed to be prescribed. The efficacy of third line anti-TB drugs are yet to be proven.

Because of the extensive use of inactive second and third line antibiotics in the treatment of MDR-TB, there is a rise in the occurrence of extensively drug-resistant

tuberculosis (XDR-TB). The XDR-TB is resistant to some key components in the second and third line drugs, including at least one fluoroquinolone and one of the injectable antibiotics (i.e. streptomycin, amikacin, capreomycin)[12]. The cure rate for the XDR-TB can be as low as 30%[13]. Even more horrifying, there are reports of totally drug-resistant TB (TDR-TB) for which there is no effective chemotherapy.

Bedaquiline is the first new agent approved for the treatment of *Mtb* since Rifampin in 1970 (Figure 1.1 c1). In 2012, the U.S Food and Drug Administration (FDA) approved Bedaquiline for use in TB treatment, but only for use in cases of severe MDR-TB or XDR-TB, which lack other treatment options. Betaquiline acts as a *Mtb* ATP synthase inhibitor, and is the first-in-class to employ this novel mechanism. There remains some controversy regarding the current clinical trial of Bedaquiline. Its manufacturer Johnson & Johnson (J&J) was able to get accelerated approval of the drug. This fast-track approval is only for cases of MDR and XDR tuberculosis. For this case, FDA's approval was based on the outcome of the sputum cultures rather than the patients' death rate. As clinical cases have accumulated, it was been shown that Bedaquiline usage increased the death rate by 5 times comparing to the control group. One significant reason for the increased death rate may be failure of treatment of the bacteria due to lack of antibacterial efficacy; there are also toxicity issues still being investigated.

1.1.2. Importance and challenges of TB drug development

TB drug development has regained scientists' attention; however, compared to the rapid development of resistance, the drug development progress is unfortunately very slow. There are at least three goals when developing new anti-tubercular drugs.

First, the standard treatment for the current drug sensitive TB needs to be simplified and shortened. The conventional “short course” regimen requires 6-month usage of RIF. Studies have shown that premature discontinuation or poor adherence in RIF usage can increase the rate of relapse and facilitate development of resistance in the disease[14]. The World Health Organization (WHO) has developed a “DOTS” (directly observed treatment, short-course) strategy to control the spread and resistance development of TB. “DOTS” has 5 main components: government commitment, smear positive detection, observed treatment, ensured drug supply and standardized reporting systems[15]. Although DOTS is considered the most effective and economical TB control strategy, it is a very expensive program that is difficult to enforce, as it requires some methods of treatment observation (i.e. hospitalizing the patients) to ensure the proper drug usage. Especially in some developing countries, patients often discontinue their therapy when their symptoms are alleviated. Thus, although DOTS and the short-course treatment being recommended, Multi-Drug Resistance problems still emerged due to poor compliance or improper treatments. Thus, the current situation calls for the development of more effective chemotherapies that can shorten and simplify the regimen and lower the treatment cost.

Second, drugs that act on new targets or through new mechanisms of action are needed for resistant TB. Since the 1990s, MDR-TB and XDR-TB has rapidly emerged, first within hospitals, later transmitted into the community. To treat MDR-TB and control the spread of resistant TB, the WHO has extended the “DOTS” strategy to “DOTS-plus”. “DOTS-plus” allows use of agents from a list of second line anti-TB drugs; even though the efficacy of these second line drugs has yet to be confirmed. So far for MDR-TB, the cure rate is 80%. For XDR-TB, however, there is not a standard strategy. Treatment often involves using of a large variety of third line antibiotics that may not be effective at all,

and the cure rate is very low (30%). It is not hard to foresee that the current inadequate treatment for MDR-TB and XDR-TB can further stimulate the rate of TB resistance development. For this reason, efforts on the development of chemotherapies working on new mechanisms are badly needed.

Third, effective cures for latent TB infection are wanted. As reported by the WHO, over 2 billion people have latent TB infection[1], and 10% of these people will develop active TB infection in their lifetime. The TB reactivation rate has significantly increased with increased cases of HIV infections, and TB infection has also become the greatest death risk for HIV patients. The high death rate for the TB co-infected HIV patients is due to lack of effective treatment. The standard 6 month regimen is problematic because, the most important chemotherapy RIF cannot be used by HIV patients undergoing antiretroviral therapy. RIF is a cytochrome P450 (CYP) 3A4 inducing agent; it has significant drug-drug interactions with many anti-HIV drugs[16]. There are efforts to identify substitutes for RIF that are compatible with anti-HIV drugs[17]. An alternative perspective to solve this issue is to develop cures for latent TB infection. So far there is no way to kill dormant bacteria that lie hidden during latent infection. Available *Mtb* drugs only attack proliferating *Mtb*, and do not kill dormant bacteria. Developing antitubercular drugs targeting dormant *Mtb* infection will help the 2 billion people with latent TB infection.

1.2. BioA as an anti-*Mtb* target

1.2.1. *Mtb* Targets of The Biotin Biosynthesis Pathway

As we have pointed out in 1.1.2, drugs that act through novel mechanisms on new targets will be vital as chemotherapeutic tools addressing current resistance problems.

For this reason, extensive work is being done in attempt to identify novel TB druggable targets. Many reported target-based screening methods are problematic and result in low success rates in identification of compounds that enter further drug development; i.e. the hits from the initial target-based high throughput screening (HTS) assays usually turn out not being validated in the later cell based assays and animal models[18]. For example, GlaxoSmithKline performed a retrospective assessment of over 70 screening campaigns in which only 5 hit compounds were advanced for further drug development[19]. To identify new targets that can be inhibited *in vivo*, whole cell screenings were conducted where targets are paired with lead compounds with whole cell activity. In 2013, Sacchittini *et al.* identified TB drug targets using a whole cell based screening assay[18]. Using this screening method, they not only identified new TB drug targets, but also identified lead compounds with whole cell activity that could be paired with the identified targets. These targets include aspartyl-tRNA synthetase, polyketide synthase Pks13, membrane transporter MmpL3, and ESX-3 type VII secretion system EccB3. The whole cell screening method gives rise to a higher rate of lead identification. An important lesson from these different screening approaches is that working on a validated target is crucial for novel anti-TB drug development.

To identify all molecular mechanisms that are crucial for the survival of *Mtb* during infection, Sassetti *et al.* mutated every gene not known to be essential of *Mtb* and discovered that every gene in the biotin synthetic pathway is essential for *Mtb* infection in a mouse model (except *bioD* for which the data is not available)[20]. This finding is meaningful for anti-TB drug development, because unlike bacteria or plants, mammalian hosts do not have enzymes for biotin cofactor synthesis[21]. Since no homologues of these enzymes exist in humans, they may be safely targeted within bacteria in a human host.

Four enzymes are involved in the biotin biosynthetic pathway (Figure 1.2). In the first step, KAPA synthase (BioF) catalyzes the decarboxylative condensation of pimeloyl-CoA with L-alanine to form 7-keto-8-aminopelargonic acid (KAPA). In the second step, the PLP-dependent DAPA synthase (BioA) catalyzes the transamination reaction to produce 7,8-diaminopelargonic acid (DAPA) from KAPA; in this reaction, *S*-adenosylmethionine (SAM) is used as the amino donor. Dethiobiotin synthetase (BioD) catalyzes the third step, the carboxylation of DAPA to produce the imidazolidin-2-one ring of dethiobiotin (DTB). Finally, biotin synthase (BioB), an iron−sulfur cluster enzyme, converts DTB into biotin by catalyzing the C-H activation and insertion of sulfur into DTB in the fourth step of biotin synthesis.

Among the four enzymes in the *Mtb* biotin biosynthesis pathway, BioA is the most well studied. BioA and BioD are the only two proteins in this pathway that have been structurally characterized[21]. Sae Woong Park et al. in 2011 published their study that further confirmed that the expression of BioA protein is essential for the acute and chronic infection of TB in a murine model[22]. In their experiment, a knockout and a knockdown of *Mtb* BioA gene models were built. The experimental results showed that without *Mtb* BioA expression, *Mtb* cells are unable to survive in vivo or in mice without a supplemental biotin source. From these experimental results, BioA is validated to be a novel effective target for both active and latent TB infection.

BioA (also known as DAPA synthase) is a Class I aminotransferase responsible for the conversion of 7-keto-8-aminopelargonic acid (KAPA) to 7,8-diaminopelargonic acid (DAPA) employing *S*-adenosyl methionine (SAM) as the amino donor. Catalysis is facilitated via a ping-pong bi-bi mechanism using PLP as a co-factor[23] that cycles

between the reduced pyridoxamine (PMP) and oxidized (PLP) states (Figure 1.3). First, SAM reacts with the internal aldimine of PLP and Lys283 to donate an amino group to the PLP, forming pyridoxamine phosphate (PMP). KAPA then reacts with the PMP-bound form of BioA, receiving the amino group to form the product DAPA. With this last step, BioA returns to its PLP bound *holo* form.

Generally, transaminases can use many endogenous amino acids as amino donors, but BioA has an unusual preference for SAM, which more often serves as a methyl donor[24]. SAM is structurally very different from the amino accepter KAPA, and how the enzyme retains specificity in the recognition of these two different substrates is only partially understood. An *Mtb* BioA X-ray structure with bound PLP co-factor was first reported by Dey *et al*. in 2008[21] (PDB-id 3BV0); the resolution and structure quality was improved by Geders *et al*. in 2011 (PDB-id 3TFT)[25]. Dey *et al*. also worked toward substrate bound *Mtb* BioA structures using X-ray crystallography. Unfortunately, they were not able to obtain *Mtb* BioA complex structure with the presence of SAM or KAPA in the active site. Using a H315R single site mutant *Mtb* BioA protein and an unreactive SAM analog sinefungin, they were able to capture a complex structure (PDB-id 3LV2)[21]. Based on the comparison of this structure to an *E. coli* BioA complex with KAPA (PDB-id 1QJ3)[26], Sacchettini and co-workers have proposed that *Mtb* BioA catalyzes the transamination by an induced fit mechanism that relies upon active site conformational changes that accommodate structurally different substrates[21]. Although conformational details of *Mtb* BioA catalysis are yet to be investigated, this induced fit mechanism suggests that substantial conformational changes must occur during the catalytic cycle. This necessary capacity for change in the active site conformation makes it challenging to design *Mtb* BioA inhibitors and predict ligand binding using computational modeling

methods, particularly based on homology models derived from structures from other species.

1.2.2 Current status of *Mtb* BioA inhibitor development

Limited efforts to identify selective inhibitors of *Mtb* BioA have been reported. Amiclenomycin (ACM), a natural product extracted from Streptomyces strains, and a simplified derivative were identified many years ago as mechanism-based inhibitors of *Mtb* BioA[23,27]. ACM showed good activity against *Mtb* cells, but failed in animal models[27] likely due to its low chemical stability[28].

More recently, mechanism-based inhibitors based on ACM (compound **1**, Figure 1.4) with improved chemical stability have been described, but the stability comes at the expense of lower potency[29]. The design of this new mechanism-based inhibitor was based on the hypothesis that by introducing a heteroatom into the 1,4-cyclohexadiene, the aromatic stabilization energy of this ring would decrease so that the compound's chemical stability against self-aromatization would increase. Biochemical evaluations and structural characterization confirmed that compound **1** is also a mechanism-based inhibitor with a two-step mechanism of inactivation (Figure 1.5). It possesses a comparable k_{inact} to ACM; however, the K_I value is 26-fold higher. This indicates that the initial binding of compound **1** with *Mtb* BioA in the active site is weak. The second-order rate constant for inactivation k_{inact}/K_I of 1 is 346 M^{-1} min^{-1}. It is a potential lead compound which can be further developed for therapeutic purposes.

For the purpose of understanding the inhibitory mechanism of compound **1**, our group produced complex X-ray crystal structures of BioA before and after the inactivation with compound **1** (PDB-ids 3TFT and 3TFU). The post-inactivation complex shows complete

cleavage of the internal aldimine and the formation of a covalent adduct, which coincides with the proposed mechanism. Structural studies with the post-inactivation complex has brought to our attention the flexibility of the BioA active site; a Trp64 side chain flipped to a new position to favor the binding of the aromatized ring in the adduct complex[25]. Nevertheless, much remains unknown about the flexibility of the active site and its capacity for binding ligands. Unanticipated conformational changes can make it difficult to predict ligand binding using the BioA *holo* structure as an initial model. A need therefore remains to better understand the conformational dynamics of BioA. More diverse chemical scaffolds that bind BioA active site can help characterize the flexibility of the active site, as well as to serve as starting points for BioA inhibitor design. A fragment-based approach can serve these purposes efficiently.

1.3. Fragment Based Inhibitor Design

1.3.1 Fragment-based drug design vs. High throughput screening

Fragment-based drug design (FBDD; also known as fragment-based lead discovery, FBLD) is a concept that has been rapidly popularized in the last decade. FBDD's main objective is to identify fragment hits (small molecule ligands) of the targets and design larger potent compounds by modification of these hits. "Fragment" is a concept to differentiate the small molecules used in the fragment screening from those often used in traditional high throughput screening (HTS). FBDD affords some distinct advantages over HTS. First, FBDD screening covers a relatively large chemical space using a small number of compounds. Unlike HTS, FBDD usually begins with a screen of several hundred or thousand compounds as combinations of fragments that cover a larger chemical space more efficiently than HTS. Second, fragment screening has a higher hit rate than traditional HTS; large complex molecules used in HTS must simultaneously

fulfill the requirements of many different sub-sites, while small a fragment molecule only need to match one. Approximately half of HTS fails because there are no initial hits discovered as starting points[30].

The disadvantages of FBDD are also not trivial. First, fragment hits often have low binding affinities to targets (high micromolar, μM or even millimolar, mM); this increases the difficulty in detecting fragment hits. Fragment hits binding affinities are often too weak to show measurable inhibition in a typical biochemical activity assay. Advanced biophysical techniques including Differential Scanning Fluorimetry (DSF), Surface Plasmon Resonance (SPR), Isothermal Titration Calorimetry (ITC) and Saturation Transfer Difference NMR (STD-NMR), are reported to be sensitive enough to detect the binding of small molecules with dissociation constants (K_D) as high as 1 mM.[31] However, to detect hits with low binding affinities, compound solutions need to be prepared at very high concentration; this requires that the fragment compounds also have very high solubility. Second, while sensitive biophysical methods can detect binding, they can provide no discimination between *relevant* binding (that which has the potential to inhibit protein biological function) from *indiscriminant* and unimportant binding. The relevance of binding to function is typically inferred only following structural characterization of the binding, mode. Structural characterization and screening must be applied together to prioritize fragment hits for further effort. The third disadvantage is also related to the low binding affinity of the fragment hits. Although it is easier to find initial hits from fragment screening than HTS, the success rate of developing the fragment hits into lead compounds is low.[31] Some fragments with very low binding affinity to a protein might not be a good starting point for further potency optimization.[32] Because fragments are by definition small molecules, and the size of a compound can potentially influence the binding affinity, a measurement of ligand efficiency (LE) is used as a comparative

measure of the value of fragment hits.[33] Ligand efficiency is defined as the binding energy per heavy atom of a ligand compound with its target protein.[34] A compound with a K_D of 10 nM and a molecular weight of 500 Da (about 38 non-hydrogen atoms) would have a LE of 0.29 kcal/mol/heavy atom. Thus, preference is often given to fragments with a ligand efficiency greater than 0.3 kcal/mol/heavy atom to start further optimization.[33]

1.3.2. Strategies for Fragment Lead Optimization

In a typical procedure of FBDD, fragment screening is first applied to identify small molecules that bind the potential drug target. This initial screen usually utilizes some biophysical technique. Initial screening is then followed by structural characterization using X-ray crystallography or NMR. Initial screening is more efficient than direct structural observation, so it can be used to reduce the number of compounds that must be examined structurally. The three dimensional visualization of fragment binding is very important because fragment molecules are small, and they usually only occupy a sub-site of the binding site with very moderate initial binding affinity. The detailed structural characterization of the fragment binding is often used to guide the further optimization. To arrive at a lead compound from fragment hits, different strategies for fragment optimization have been developed. These strategies may generally be grouped into two categories: “fragment linking” and “fragment growing”. The fragment linking strategy can be used when two fragment hits are discovered to bind in adjacent sub-sites of the protein. Ideally, the two fragments can be combined to form a larger molecule with higher affinity by contriving a chemical linker that joins the two fragments without perturbing the binding of each; in the real world, fragment linking is a very challenging strategy. There are several requirements for the fragments that are linked together. First, the distance between the two binding fragments should be proper: if the fragments are

too close together, there might not be space to place the linker, but if the fragments are too far away, the long linker might result in too large a molecule. Second, the incorporation of the linker cannot affect any important interactions of the fragments with the protein or create strain that changes the fragment binding conformation. Third, the linker itself should not introduce unfavorable interactions with the target. These requirements make it difficult to successfully employ a linking strategy[31,35]. For the reasons above, there are few successful examples of lead compound generation by fragment linking since it was first introduced by Shuker *et al.* in 1996[36]. There are several successful examples of fragment linking involving protease targets[31,37-39], as proteases often have extended binding sites. In one example, Howard *et al.* discovered thrombin inhibitors with nanomolar binding affinity by fragment linking[40]. They used X-ray crystallographic screening and identified fragments that bind to the S1 pocket and S2-S4 pockets respectively. Fragment linking has also been successfully applied to acetylcholinesterase, which has a deep, gorge-like active site[41]. In 2004, Harry *et al.* designed an inhibitor of acetylcholinesterase by linking fragments that bind in the active site and an adjacent cation binding site[41].

Compared to fragment linking, fragment growing is more often used for lead discovery. In a fragment growing strategy, a stepwise extension of the fragment is often performed to gain additional interactions with the protein target[42-44]. In some real FBDD cases, the distinction between “linking” and “growing” can be unclear. In the initial fragment screening, when there are multiple fragments binding in the same binding site but making different interactions with the protein, one commonly used strategy is to merge the molecular features of different fragments to devise compounds with more interactions and better binding affinity[35]. Fragment binding information can also be used

for the optimization of lead compounds by scaffold hopping, where new ligands are created by an exchange of chemical features between existing ligands.[45]

1.3.3. Important experimental components for FBDD

A fragment-based drug discovery approach necessarily begins with empirical fragment screening conducted to identify a collection of small molecule fragments that bind to the target. To successfully complete a fragment screen, several experimental components must be in place including: 1) A good "fragment library"; 2) a sensitive and efficient biophysical method to detect compound binding, and 3) a means to structurally characterize ligand binding. We will elaborate on these points one by one.

1.3.3.1. Fragment libraries

The composition of the fragment library is crucial for success in fragment screening. In general, there are two types of fragment libraries: diverse libraries and focused libraries. Diverse libraries are designed to efficiently cover a large chemical space for more general screening purposes. Diverse fragment libraries are often used for the initial exploration of possible binding features of the targets. A successful fragment screen relies heavily on the quality of the fragment library. An ideal fragment library should possess the following features. First, the compounds in the library need to have superior physicochemical properties to allow room for further development into molecules with suitable ADME attributes.[46] The "Rule of 3" is a widely accepted criteria for fragment physicochemical properties (the molecular weight is less than 300 Da; there are no more than three hydrogen bonding donors or acceptors in the molecule; the partition coefficient (clogP) is to be less than three; and the number of rotatable bonds should be no more than 3).[35,47] Second, a general purpose fragment library needs to cover a

considerably large chemical space.[48] Clustering software is available to assist in design of highly diverse libraries.[49,50] Third, for the purpose of drug design, it is often beneficial to start with pharmacophore enriched molecules.[48]

A target-focused library is often specially designed around a certain chemical scaffold or a key pharmacophore and used for targeting a particular protein or protein family. The design of a focused library can be assisted by computational tools.[51] Libraries can either be designed based on the knowledge of the targets[52] or from ligand similarity search using fingerprints of molecules[53]. There are multiple examples of successful implementation of a focused library in drug discovery. In 2005, Grasberger *et al.* screened a large benzodiapepine library (> 300,000 compounds; which is used as a collection of peptide mimetics for probing possible protein-protein interaction sites) and discovered a novel series of benzodiazepinones as HDM2−p53 interaction antagonists.[54] In a nice example from 2002, Liebeschuetz *et al* demonstrated use of a very small but hit-rich library of benzylamido analogs that were already known to preferentially bind in the S1 subsite of Factor Xa.[55]

For the *Mtb* BioA case, because there is limited knowledge of structural features of bound ligands, a general purpose library is appropriate. We used a diversity core set of Maybridge Ro3 library (1000 compounds) for screening purposes. The fragments in this library have many good features. First, they follow the “Rule of 3”. Second, they have assured solubility in both DMSO (200mM) and PBS buffer (1mM). Third, the library is computationally engineered to have diversity; it is pharmacophore rich but includes structures with low complexity.[56]

1.3.3.2. Biophysical screening tools used in FBDD

Since small molecular fragments may not be expected to possess measurable activity in biochemical enzyme assays, fragment screening must be conducted with sensitive but efficient biophysical methods that can be applied to detect small molecule *binding*. Several methods have been widely used in fragment screening applications.

Differential Scanning Fluorimetry (DSF) can be used to detect small molecule binding to proteins and has been used in fragment screening.[57-59] This is based on a fact first introduced by Koshland in 1958 [60], that binding by a small ligand may cause conformational changes of a protein, even regardless of enzymatic activity. Changes in protein stability that accompany these conformational changes can shift the midpoint of the heat-induced transition from the folded to unfolded state (the "melting temperature" or T_m). Unfolding can be monitored by DSF if a fluorescent dye sensitive to the appearance of unfolded protein is included in the solution along with the potential ligand. DSF is a medium-throughput assay. By using a real time-PCR (RT-PCR) machine, up to 384 fragments can be screened simultaneously in one plate[61].

There are varieties of choices for the fluorescent dye used in DSF. Generally speaking, there are three features of the dyes: 1) the intrinsic fluorescence of the dye must be quenched by water in solution; 2) the dye binds nonspecifically with hydrophobic residues; 3) the fluorescent property of the dye changes significantly upon binding with the hydrophobic residues. Because of the usage of the dye, DSF cannot be considered a completely label free method. However, the protein is not pre-labeled before the experiment, and the possible effects of the dye on the protein and small ligands can be characterized using several control experiments. DSF can be used to evaluate binding of small ligands with a K_D value of 10^{-9} to 1 M. It provides a huge sensitivity range compared to other techniques. DSF also consumes a relatively small amount of protein

(about 0.4 μg/experiment). DSF does not provide information about binding sites or stoichiometry, so a structural characterization of hits is a necessary to confirm relevance. DSF can also produce both false positive and false negative results. Although ligand binding may induce either a positive or negative T_m shift, most investigators have typically focused only on ligands that stabilize the protein to unfolding.[62] We will elaborate on this point more in later discussion.

Saturation Transfer Difference NMR (STD-NMR) is a ligand based NMR technique. It is one of the most popular techniques in use for fragment screening[38,63], and was in fact the first biophysical method used in a fragment screening paradigm.[36] STD-NMR utilizes the fact that a small molecule has a much faster relaxation time than a large molecule, and that the spin excitation on a large molecule can be transferred to a bound small molecule and used to detect ligand binding. In a typical STD-NMR experiment, two NMR signals are given: on-protein resonance frequency and off-protein resonance frequency. By subtracting the two NMR spectra signals, the signal of the bound small ligand will be detected. Rapidly evolving NMR capabilities (higher resolution, auto sampling systems etc.) has allowed STD-NMR screening to be a high-throughput screening technique. First, multiple small molecules can be screened in one single experiment as long as their spectra are differentiable. Second, auto sampling systems and shortened scanning time for each experiment make STD-NMR quite efficient. Another great advantage of STD-NMR is that it is a completely label free experiment that reflects the solution state of the protein; the experiment can be performed under a steady temperature of choice to best mimic the physiological conditions. The detection range of STD-NMR is 10^{-8}-10^{-3}M: binding affinities that are too tight or too loose can cause problems based on the underlying principle of this method. STD-NMR consumes a relatively high amount of protein in a single experiment (up to 50-75 μg)[64]. Like DSF screening, STD-NMR also

cannot provide information where compounds bind. However, there is one unique piece of information STD-NMR can provide that is not provided by other biophysical screening techniques, the epitope maps. The strength of the saturation transfer signal of a particular atom in the small molecule is negatively correlated to its distance from the protein: the closer the atom is towards the protein, the stronger its saturation transfer signal is. Using this feature, relative distances of the detected atoms from the protein can be mapped out. This map can be informative as it shows which part of the small molecule is more important for binding.

Surface Plasmon Resonance (SPR) is another popular technique for fragment screening.[65-67] In SPR, the protein is immobilized on a metal surface of a specially designed chip, and the ligands flow past the surface during the experiment. Ligand binding can cause changes in reflectivity properties of this surface. These changes are related to some properties of the surface, as well as the mass of the protein and ligands. With the current techniques SPR is not a high throughput screening method, but it can be used to screen thousands of compounds needed for fragment screening. SPR consumes very low amount of protein, because once the protein is immobilized on a surface of a chip, the chip can be used for screening of many different ligands. There are in general two different methods for protein immobilization: the covalent coupling chemistry and unidirectional immobilization. Covalent coupling chemistry uses the reactive groups on the surface of the protein (i.e. amine, thiol etc.) to form covalent linkage with the surface. The reacting group is randomized so the orientation of the protein is also randomized. Unidirectional immobilization tethers the protein to the surface in a certain orientation. It utilizes some site-specific immobilization tool such as antibody and biotinylation to build a more homogenous protein surface. Neither of these immobilization methods can retain the proteins' natural behaviors, and this might cause

some false screening results. The detection range of SPR experiment is about 10^{-3}-10^{-9} M, and it is dependent on the solubility of the ligands. SPR can be used to study the kinetics of the ligand binding. It gives a direct measurement of the association rate constant k_{on} and dissociation rate constant k_{off}, SPR is also widely used to characterize the ligand binding affinity, especially for ligands with low binding affinities.

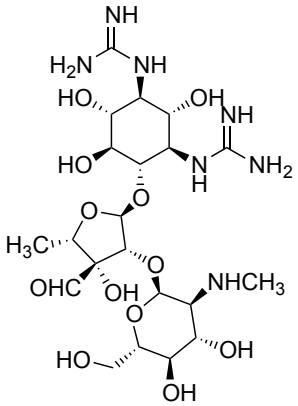
1.3.3.3. Structural validation tools for fragment screening

NMR is often used to validate fragment binding. Target-based approaches were developed first to characterize ligand bindings.[68] This approach compares 2 dimensional NMR spectrum of the protein alone with the spectrum of the protein in presence of the ligand. Changes in protein chemical shifts indicate binding. Provided chemical shift assignments have been completed for the protein structure, the binding site can be located by identifying protein residues with chemical shift changes. This target-based NMR technique has several advantages. First, this is a very robust method to validate ligand binding and it can provide some structural information about binding mode. Second, using NMR, the experiment can be designed under solution conditions that best mimic the proteins' endogenous conditions. Third, it is almost the only way to analyze structures of proteins that are difficult or impossible to crystalize. There are also caveats of this method. First, this experiment consumes a large amount of protein.[69] In some cases hundreds of milligrams are used for one experiment.[70] Second, this approach has a limit for the protein size (30-40 kDa)[68]. This is because ^{15}N enriched protein needs to be expressed, and chemical shift assignment for the entire protein structure is required. To overcome these technical problems, ligand based NMR approaches have emerged recently. These approaches can be used to acquire ligand NMR spectrum and detect ligand binding by observing ligand chemical shift changes rather than protein; however, they do not provide structural data.

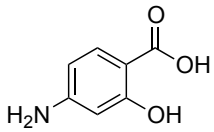
X-ray crystallography may be used to validate the binding of fragment hits discovered from the screening assays. Getting a complex structure model with X-ray crystallography is considered to be a “gold-standard” for the validation of fragment binding. A crystal structure can provide a three-dimensional binding mode of the ligand with the protein. It can be used to distinguish specific binding from nonspecific binding; it can also show if a ligand is binding in the active site or in some allosteric or irrelevant sites. However, there are potential liabilities of crystallography: diffraction analysis can only be conducted following a successful search for crystallization conditions. Protein crystallization solutions are also highly concentrated, or may require high ionic strength precipitants that can often cause problems with the co-crystallization of compounds with poor solubility. For this project, because conditions were known for the production of high quality crystals of BioA, X-ray crystallography was selected as the method for structural validation and structure-based fragment optimization. Information from X-ray crystallography provides invaluable information to guide the molecular design of new inhibitors.

Figure1.1. Structures of some important TB drugs

a. First antitubercular drugs

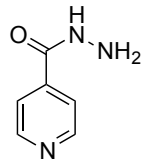


a1. Streptomycin (STR)

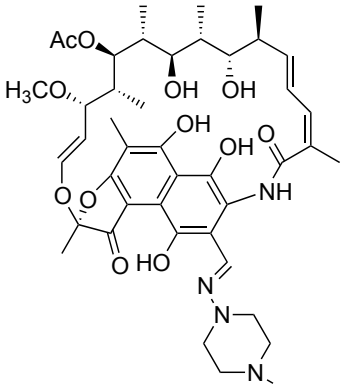


a2. *para*-Aminosalicyclic acid (PAS)

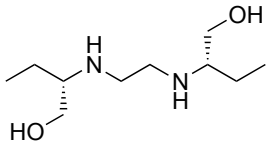
b. Currently used firstline antitubercular drugs



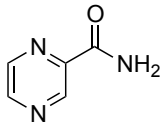
b1. Isoniazid (INH)



b2. Rifampicin (RIF)

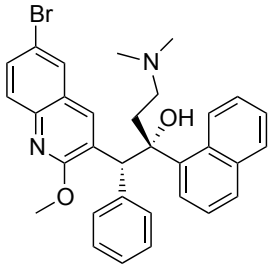


b3. Ethambutol (EMB)



b4. Pyrazinamide (PZA)

c. Recently approved antitubercular drug



c1. Bedaquiline

Figure 1.2. *Mtb* BioA biotin biosynthesis pathway.

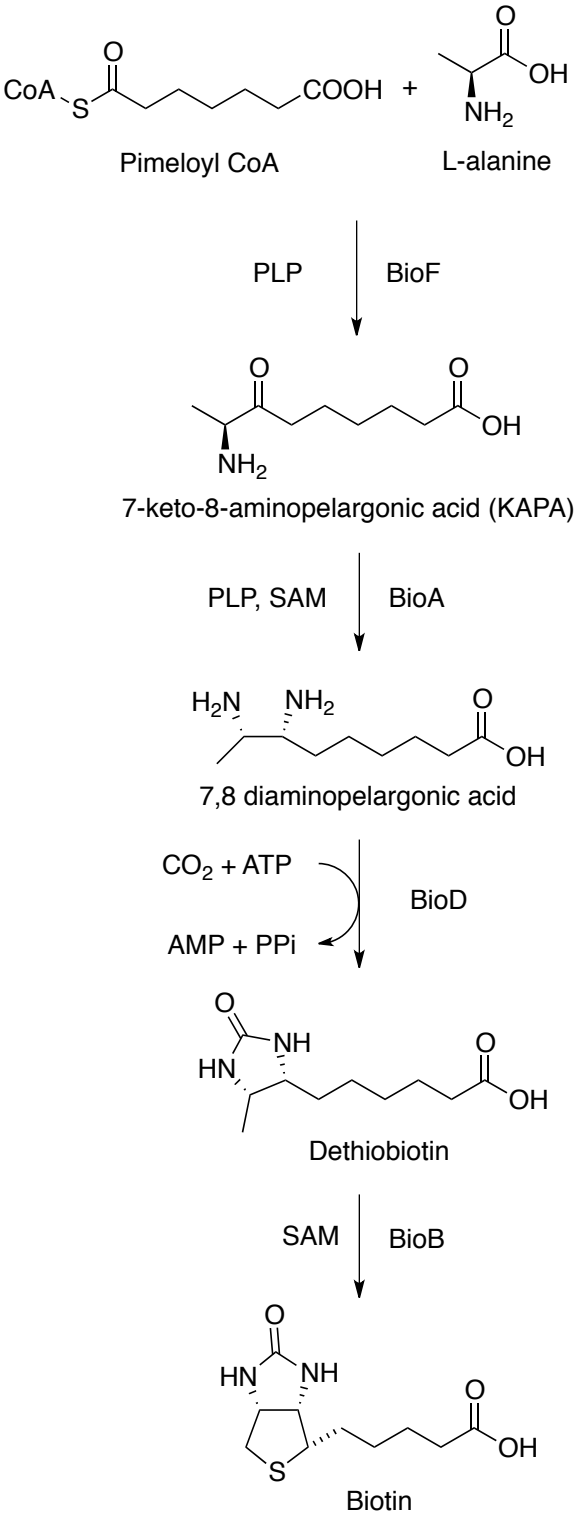


Figure 1.3. Catalysis of DAPA synthesis by BioA.

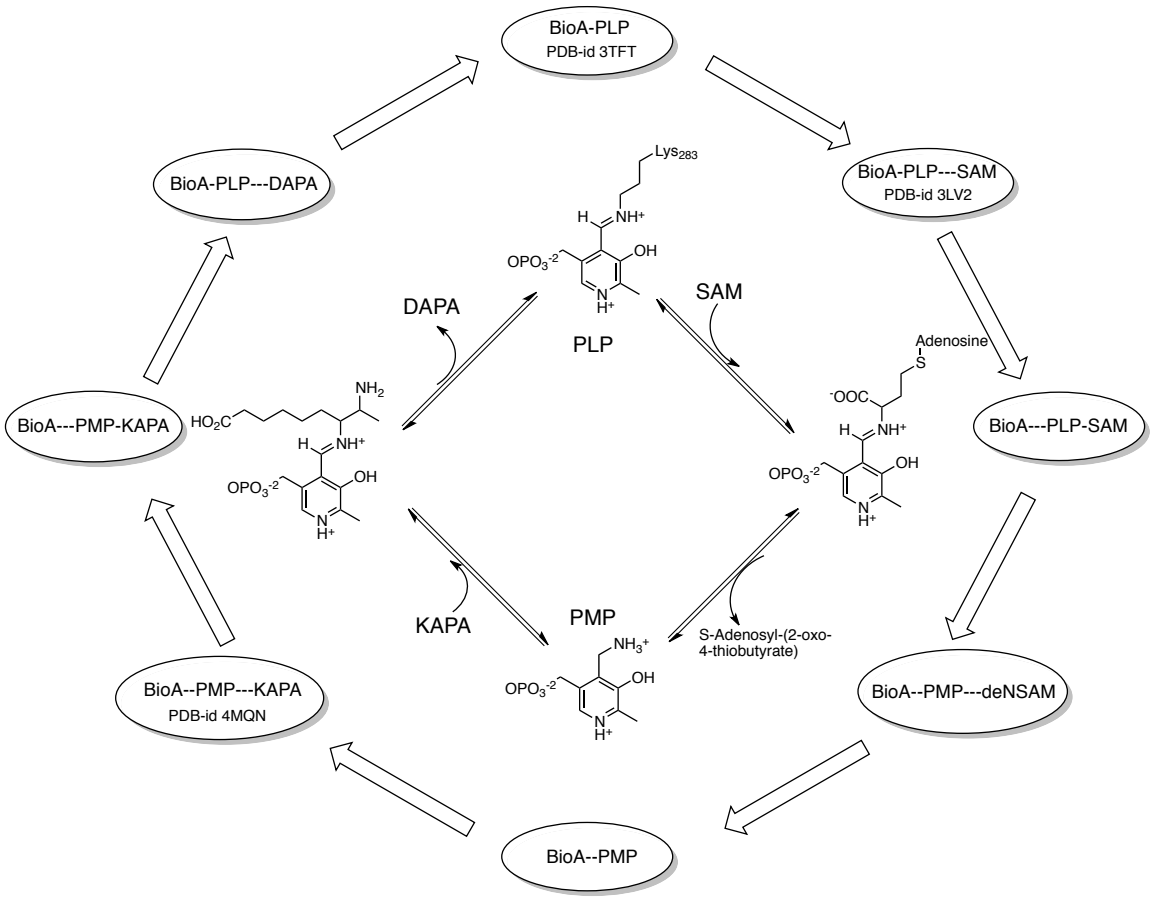
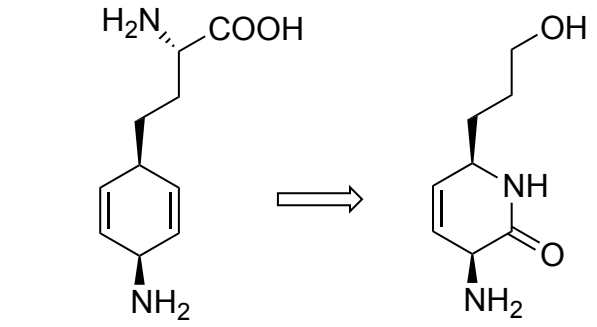
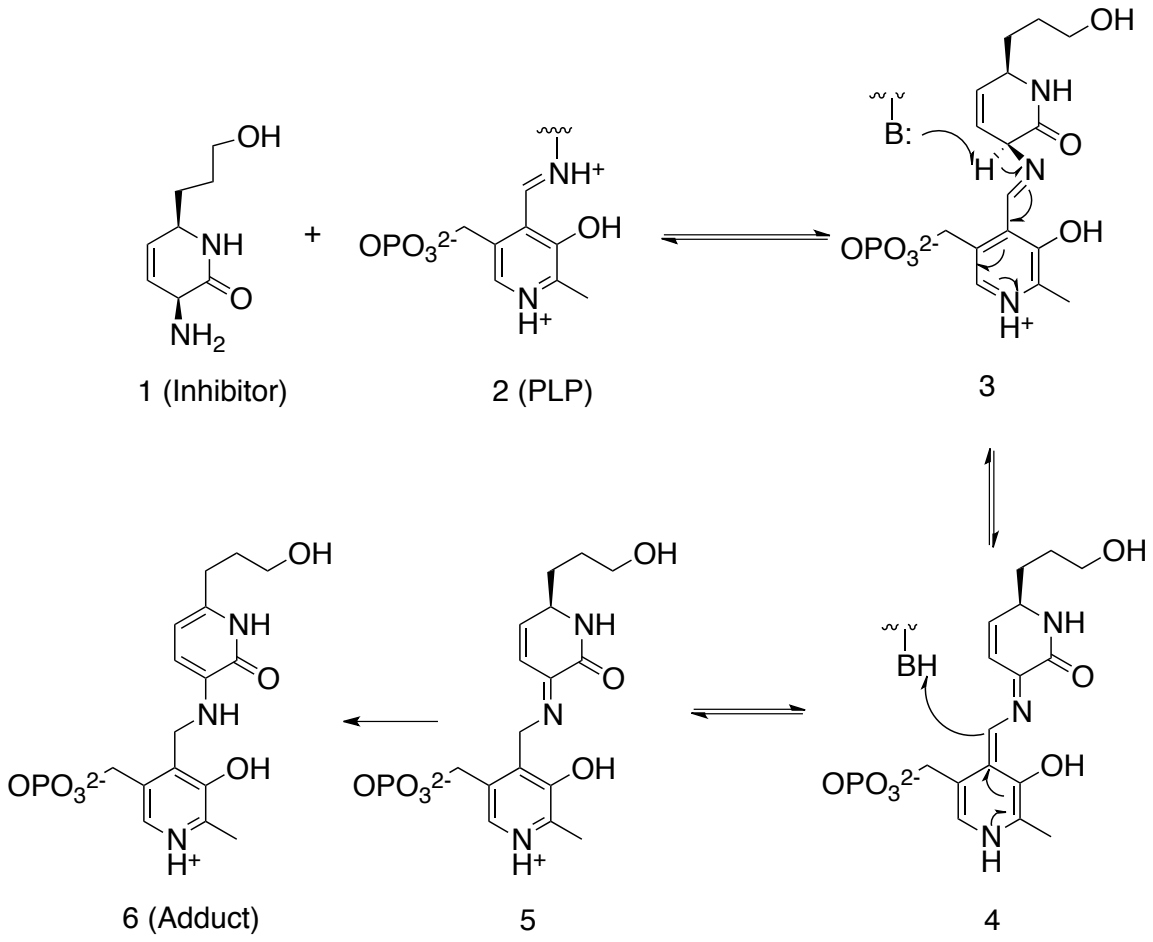


Figure 1.4. Recent Mechanism-Based Inhibitor from Natural Product Modification.



Amiclenomycin (ACM) **1**

Figure 1.5. Inactivation mechanism of compound 1 by adduct formation with the BioA PLP co-factor.



Chapter II: Crystallographic studies of *Mtb* BioA with substrate KAPA

Acknowledgement: Some material included in this chapter is excerpted from previously published work [Dai, R. et al, 2014. "Inhibition of Mycobacterium tuberculosis Transaminase BioA by Aryl Hydrazines and Hydrazides ". ChemBioChem. 15(4):575-586]. This includes sections 2.2.4, 2.3.2 and figures therein. Ran Dai and Barry Finzel contributed equally in writing this published work, with assistance from other authors. Permission to use this copywrited material has been granted by John Wiley & Sons.

2.1 Introduction

2.1.1. Review of previous *Mtb* BioA production and crystallization efforts

In 1998, Cole *et al*. published their work describing the complete genome for *Mycobacterium tuberculosis*.[71] As we have discussed in Chapter 1, enzymes involved in the biotin synthesis pathway were identified as promising as drug targets[20] and BioA, the enzyme that catalyzes the second step, is of particular interest since it is both structurally[21] and functionally[72] well characterized.

In 2010, the first effort toward overexpression, purification and crystallization of *Mtb* BioA was reported by Dey *et al*.[21] For overexpression, they constructed a plasmid by cloning *bioA*(Rv1568) into a pET28b vector with the TEV site using the *Nde*I and *Hin*dIII sites. IPTG was used to induce the BioA overexpression. Purification was accomplished with a Histrap nickel column using a tris(hydroxymethyl)aminomethane (Tris) buffer system. Crystallization of BioA was achieved under a PEG-based condition using Tris as the pH

buffer (10% PEG 8000, 0.1 M Tris buffer pH 7.0, and 0.1 M $MgCl_2$). Diffraction data was collected to 2.2 Å resolution using a home X-ray source[21].

In 2011, Wilson *et al*. published an alternative plasmid construct for BioA overexpression[73]. They first cloned *bioA*(Rv1568) into a pET28b vector including a TEV site using the *Nde*I and *Hin*dIII sites. Then they cloned *bioA* along with the RBS and His tag from pET28b into a pUC18 vector to yield a new plasmid pCDD126; pCDD126 is controlled by a *lac* promoter so the protein expression does not require induction by IPTG.

In 2012, Geders *et al.* published an improved protocol for BioA purification and crystallization[74]. They used differential scanning fluorimetry (DSF) to monitor the homogeneity of the protein and discovered that by using 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) as the pH buffer rather than Tris, and *tris*(2-carboxyethyl)phosphine (TCEP) as reducing reagent instead of dithiothreitol (DTT), the homogeneity of the BioA protein can be significantly increased. Removal of Tris as the buffer in purification increased the purity of BioA final product, because Tris can potentially react with the PLP co-factor of BioA.[74] In their purification protocol, an additional size exclusion column (HiPrep 26/60 Sephacryl S200 HR column from GE healthcare) was used after the Histrap nickel column for extra purification. It was also reported that by using HEPES as the pH buffer in crystallization, the quality of the BioA *holo* crystals could be enhanced so that diffraction data with higher resolution could be obtained [25,74]. Methods for *Mtb* BioA expression, purification and crystallization are compared in Table 2.1.

In our research, the expression system developed by Wilson *et al.* and the optimized BioA holo protein purification and crystallization protocols were used to achieve the best protein quality and crystallography results. Detailed procedures are described in section 2.2.

Table 2.1. Methods for BioA expression, purification and crystallization.

	Dey *et al*. (2010)	Wilson *et al*. (2011)	Geders *et al*. 2012
Expression			
Gene amplified	*bioA* (Rv1568) from *Mtb* H37Rv genomic DNA	*bioA* (Rv1568) from *Mtb* H37Rv genomic DNA	*bioA* (Rv1568) from *Mtb* H37Rv genomic DNA
Vector	pET28b with TEV site	pUC18	pUC18
His tag	Yes, from pET28b	Yes, from pET28b	Yes, from pET28b
Promoter	T7 promoter	*lac* promoter	*lac* promoter
Media	Luria–Bertani (LB) cultures	Luria–Bertani (LB) cultures	Terrific Broth (TB)
Production and purification technique for cell lysing	French press	Sonication	Sonication
Columns used	1. Hitrap nickel column (GE Healthcare); 2. dialysis for buffer exchange	1.A gravity flow column withNi-NTA agarose resin; 2.PD-10 column (GE Healthcare) for buffer exchange	1. HisTrap HP Ni-NTA column (GE Healthcare); 2. HiPrep 26/60 Sephacryl S200 HR column (GE Healthcare)
pH buffer for storage	Tris	Tris	HEPES
Reducing reagent	DTT	2-Mercaptoethanol	TCEP
Crystallization condition	10% PEG 8000, 100 mM Tris buffer (pH 7.0), and 100 mM $MgCl_2$	Not applicable	8%-12% PEG 8000, 100 mM $MgCl_2$, 100 mM HEPES pH 7.5

2.1.2. Structural studies of the *Mtb* BioA catalytic cycle

As discussed in Chapter 1, BioA catalyzes amino transfer reaction via an induced-fit mechanism [26 28]. Detailed structural descriptions of the conformations of the enzyme intermediates in the catalytic cycle are very limited. Dey *et al*. have described a complex structure of a mutated construct of BioA with a SAM analogue sinefungin in the active site (pdb id 3LV2)[21]. However, no KAPA bound *Mtb* BioA structures were available to study the detailed conformational changes in the active site in the catalytic cycle. A KAPA bound *E.coli* BioA model (pdb id 1QJ3)[26] and a KAPA bound *B. subtilis* BioA model (pdb id 3DU4) were used to predict *Mtb* BioA KAPA binding. However, *E.coli* BioA only has 48% sequence identity with *Mtb* BioA[75] and *B. subtilis* BioA model has only 32% identity with *Mtb* BioA[21]. To better understand the detailed active site conformation of *Mtb* BioA when KAPA is bound, we generated and studied the *Mtb* BioA KAPA complex using X-ray crystallography (pdb id 4CXQ)[62].

2.2. *Mtb* BioA expression, purification and crystallization

2.2.1. *Mtb* BioA expression and purification

The experimental procedure for *Mtb* BioA expression and purification is as described by Geders *et al*. [74]

2.2.2. BioA holo protein homogeneity assessment using DSF

Differential scanning fluorimetry was performed using a protocol established by Todd Geders [74]. A total volume of 40μL solution was used for one DSF experiment. BioA was diluted on ice to a final concentration of 0.1 mg/ml in a solution containing 25 mM HEPES pH 7.5, 50 mM NaCl and 5X SYPRO Orange (Life Technologies). To generate the DSF melting curve, fluorescence was measured using the FRET channel of the CFX96 between 298 and 368 K with 30 s incubation per 1°C temperature increase.

2.2.3. *Mtb* BioA *holo* protein Crystallization

The crystallization conditions used were as described by Geders *et al.*[21 74]. These BioA *holo* protein crystals were used for crystal soaking experiments for the fragments in chapter III, and to prepare seeds for all crystallization experiments in this thesis. The seeds were prepared by crushing small crystals from previous vapor diffusion drops. To prepare BioA seeds, a 2 μl drop of crystals was diluted in 100 μl reservoir solution (15% PEG 8000, 100 mM HEPES pH 7.5, 100 mM $MgCl_2$) and crushed thoroughly; the seed concentration of this solution was defined as 1:1. By consecutive 10 fold d ilution of the 1:1 seed solution, 1:10, 1:100, 1:1000 and 1:10000 seed solutions were made. From our experience, 1:1000 seed concentration should be used for optimal crystallization results. Crystals appear in the drop within 24 hours and grow to their full size in 72 hours. BioA *holo* crystals were protected by transferring to a cryo solution (15% PEG400, 15% PEG 8000, 100 mM HEPES pH 7.5, 100 mM $MgCl_2$) and then they were flash frozen in liquid nitrogen. Because crystal nucleation and growth always occurs with uncertainty, crystallization is always conducted as a series of experiments in which precipitant and seed solution contributions are varied systematically to surround the anticipated optimum conditions. Detailed crystallization conditions are listed in Table 2.2. The BioA

holo crystals prepared in this way were used for soaking with compounds to obtain complex structures.

Table 2.2. Example of *Mtb* BioA crystallization conditions.

BioA (A280 = 13, equals 10 mg/mL) in 25 mM HEPES pH 7.5, 50 mM NaCl, 1 mM EDTA, 0.1 mM TCEP frozen at -80°C since purification and stored in Box X14

1.5µL reservoir + 2µL BioA (10 mg/ml) + 0.5 µL seeds — RT setup, 20° incubation

SEEDS stored in 15% PEG 8K 100mM HEPES7.5 100mM MgCl2 freshly prep

Stock solution	SEEDS	9.0	10.0	11.0	12.0	13.0	14.0	Final conc
50% PEG 8k	1:1,000	180	200	220	240	260	280	9-14%
1M HEPES pH 7.5		100	100	100	100	100	100	100mM
1M MgCl2		100	100	100	100	100	100	100 mM
water		620	600	580	560	540	520	
50% PEG 8k	1:1,000	180	200	220	240	260	280	9-14%
1M HEPES pH 7.5		100	100	100	100	100	100	100mM
1M MgCl2		100	100	100	100	100	100	100 mM
water		620	600	580	560	540	520	
50% PEG 8k	1:5,000	180	200	220	240	260	280	9-14%
1M HEPES pH 7.5		100	100	100	100	100	100	100mM
1M MgCl2		100	100	100	100	100	100	100 mM
water		620	600	580	560	540	520	
50% PEG 8k	1:5,000	180	200	220	240	260	280	9-14%
1M HEPES pH 7.5		100	100	100	100	100	100	100mM
1M MgCl2		100	100	100	100	100	100	100 mM
water		620	600	580	560	540	520	

% PEG 8000

A 9%-14% (w/v) gradient of PEG concentration and two different seed concentrations (1:1000 and 1:5000) were tested.

2.2.4. The BioA KAPA complex structure determination

A co-crystallization technique was used to obtain the BioA KAPA complex structure. Protein solution (10 mg/mL in 25 mM HEPES pH 7.5, 50 mM NaCl, 0.1 mM TCEP) was mixed with reservoir solution containing KAPA (9–18% PEG 8000, 100 mM HEPES pH 7.5, 100 mM $MgCl_2$, and 5 mM KAPA) and a 1:1000 seed solution in a 4:3:1 ratio. Crystals appear in the drop within 24 hours and grow to their full size in 72 hours. BioA-KAPA co-crystals were protected by transferring to a cryo solution (15% PEG400, 15% PEG 8000, 100 mM HEPES pH 7.5, 100 mM $MgCl_2$, and 5 mM compound) and then flash frozen in liquid nitrogen.

The diffraction data was collected at 100 K using synchrotron radiation with a Dectris Pilatus 6M Pixel Detector on beamline 17-ID (IMCA-CAT) at APS, Chicago, United States. The data were processed, integrated, and scaled with XDS[76] and SCALA using the autoPROC scripts available at IMCA-CAT. Summary data collection and processing statistics are given in Table 2.3. The structures were solved by molecular replacement using Phaser[77] in the CCP4 package[78] using atomic coordinates from PDB code 3TFT as a search model.[25] Refinement and model building was done using REFMAC5[79] and coot[80]. The figures were prepared with PyMOL (The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC.). Structures were superimposed for analysis and displayed using the shared BioA-PLP overlay method of the DrugSite server[81]. Atomic coordinates and diffraction data have been deposited in the Protein Data Bank with accession code 4CXQ.

Table 2.3. Summary of Crystallographic Data.

PDB code	4CXQ
Ligand	KAPA
Data collection site	IMCA-CAT
Detector	DECTRIS PILATUS 6M
Wavelength (Å)	1.000
Space group	$P2_12_12_1$
Cell dimensions	
a, b, c (Å)	62.553, 66.149, 205.294
α,β,γ (°)	90, 90, 90
molecules per ASU	2
Resolution (Å)	55.60-1.80(1.806-1.800)
R_{merge}	0.075(0.432)
I/σ_I	16.6(4.0)
Completeness	98.9% (98.3%)
Multiplicity	6.5(6.5)
No. observations	515130
No. unique reflections	79062
Refinement	
Resolution (Å)	102.65-1.80
R_{work}/R_{free}	16.8/20.0
No. atoms	7184
No. water	527
No. ligand molecule	2
No. PLP molecule	2
No. other molecule	5
Ramachandran plot	
Favored	765 (95.6%)
Allowed	29 (3.6%)
Disallowed	6 (0.8%)
R.m.s deviations	
Bond lengths (Å)	0.009
Bond angles (°)	1.24

2.3 Results and discussion

2.3.1. Optimization of BioA purification and crystallization procedures

Higher quality and yield of *holo* BioA was obtained by using the optimized purification protocol by Geders *et al* (2012): the yield of the protein was increased to 10-15 milligrams per liter of cell culture. Geders has repeated a previous protocol by Dey *et al.* (2010) and only 1.5 milligrams per liter of cell culture was obtained. Wilson *et al.* also reported a 1.5 milligrams per liter of cell culture yield using his expression and purification method[73]. In these previous protocols, an anime containing buffer 2-amino-2-(hydroxymethyl)-1,3-propanediol (Tris) and thiol containing reducing agent dithiothreitol (DTT) was used. Also, no supplemental PLP was added to the final protein solution to ensure PLP saturation. *Holo* BioA protein has its PLP cofactor covalently linked with the terminal amino group in Lysine 283 in the protein active site, forming a Schiff base. This linkage can be broken when extraneous reactive compounds (amine or thiol compounds) come to compete. In Geders' optimized purification, the reactive components were removed from the purification buffers, and extra amount of PLP was supplemented to ensure protein homogeneity during the purification procedure. These modifications greatly increased the yield and quality of the active form of the protein, and helped to increase the quality of the crystals formed in crystallization[74].

2.3.2. The KAPA complex structure

As previously described, *Mtb* BioA is a functional homodimer with two catalytic sites that lie about 18 Å apart. Each active site is comprised of loops contributed by both polypeptide chains, including residues Pro24-Ser34, Ser62-Ala67, Arg156-Asp160, His171-Arg181, Gln224-Gly228, Arg400-Arg403, Met87′-His97′, and Ala307′-Asn322′ (Figure 2.1A). Residues tagged with a prime originate in the alternate chain. Ser125,

Asp254, Lys-283 and Thr-318′ all make specific interactions with the PLP, are strictly conserved in all Class I transaminases, and adopt a conserved conformation in all BioA structures reported to date. While residues 25–33 are disordered in the original BioA structure reported by Dey *et al.* (PDB-id 3BV0),[21] Tyr25 is well-ordered in the structure reported as a "pre-reaction" conformation in our study of an irreversible inhibitor of BioA,[25] and the side chain hydroxyl of the tyrosine lies H-bonded to Asp-160 and poised to interact with substrates. This structure (PDB-id 3TFT) will stand to represent the holo enzyme (PLP-bound resting state) conformation in the remainder of this discussion. In this state, the PLP is covalently bound to the side chain amine of Lys283, a well-known feature of transaminases (Figure 2.1A).

To provide a more reliable picture of different conformational states promoted by substrate binding, we determined the structure of the KAPA-bound enzyme at 1.8 Å resolution (Table 2.3). At this resolution we can confirm from the carbon hybridization and stereochemistry that, as expected, KAPA is bound , and not the product DAPA (Figure 2.1C). KAPA binds to the *holo* enzyme much as predicted, and induces no significant conformational changes to binding site residues, except for an inward shift of the guanidinium of Arg400, which moves in to pair against the KAPA carboxylate and cap the binding pocket (Figure 2.1D). Such a shift had been predicted[21] based on similarity to the *E. coli* enzyme structure.[26] The KAPA amino group does not displace the lysine bond to the PLP, but instead sits between the Tyr25 OH and the PLP phosphate, making strong hydrogen bonds to both. Other H-bonds to Tyr-157 OH and the carbonyl O of Gly316′ complete the tetrahedral H-bonding around the protonated amine (Gly316′ is contributed by the alternate chain of the BioA homodimer). The ketone O of KAPA makes no hydrogen bonds, but is positioned just 3.4 Å from the Nz of Lys283, which

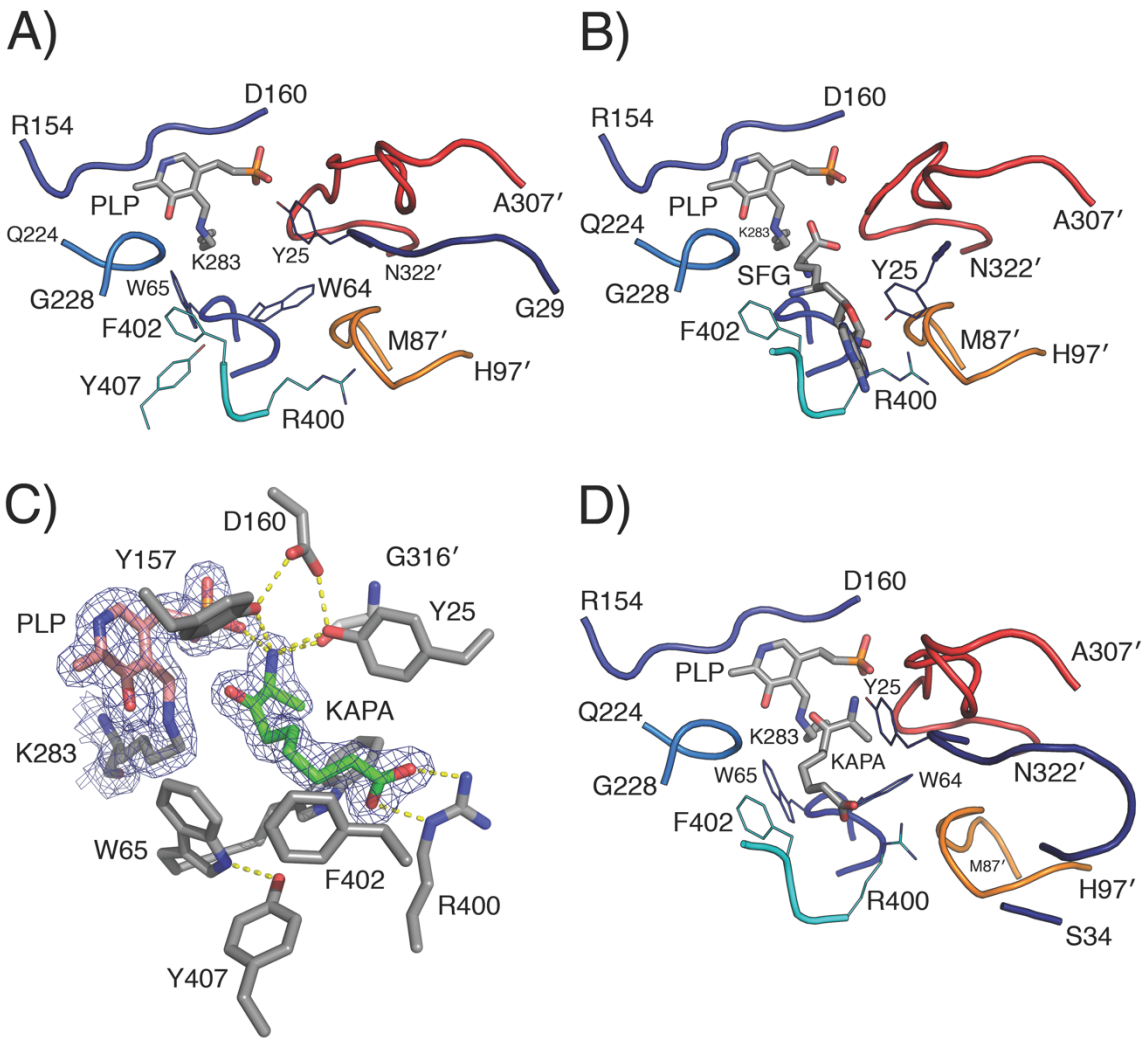
remains covalently bound to the PLP. The close packing of the β-methylene group against the aromatic ring of the Trp64 indole adds a strong hydrophobic contact to enhance binding.

Structures of *Mtb* BioA with substrates SAM and KAPA have not been previously reported, but detailed homology-based conceptions of substrate binding have been proposed based on studies with substrate surrogates (e.g., SAM analog sinefungin; PDB-id 3LV2), and homologous proteins (e.g., *Bacillus subtillis* BioA co-structure with KAPA, PDB-id 3DU4).[21] Despite the fact that the two substrates are quite different, both appear likely to occupy the same binding site, comprising a largely hydrophobic pocket between aromatic side chains of Trp64, Trp65, Tyr25 and Phe402. The adenosine of SAM likely extends further (Figure 2.1B), but apart from the side chain of Arg400, the specific groups that interact with the nucleobase are not convincingly known because of disorder observed in surrounding loops (residues 23-34, and 309-317) in many of the homolog and BioA crystal structures. This flexibility may be an important characteristic of an enzyme that must adapt to diverse substrates during the catalysis of transamination.

Dey *et al.* have reported a *B. subtilis* BioA-sinefungin complex.[21] In their report, Tyr25 is observed in a different conformation with the side chain rotated 180° to contact the adenosine base (Fig 2.3B). They conclude that π-stacking between the adenosine and this tyrosine or the phenylalanine found in other bacterial species is important feature stabilizing SAM binding in *Mtb* BioA. While we also have no specific data to contradict this hypothesis, we do not find their arguments compelling. The geometry for effective π -stacking in the sinefungin complex structure is poor, as the phenol ring of Tyr25 is tilted 45° from the plane of the adenosine; the Tyr25 OH makes the closest contact to

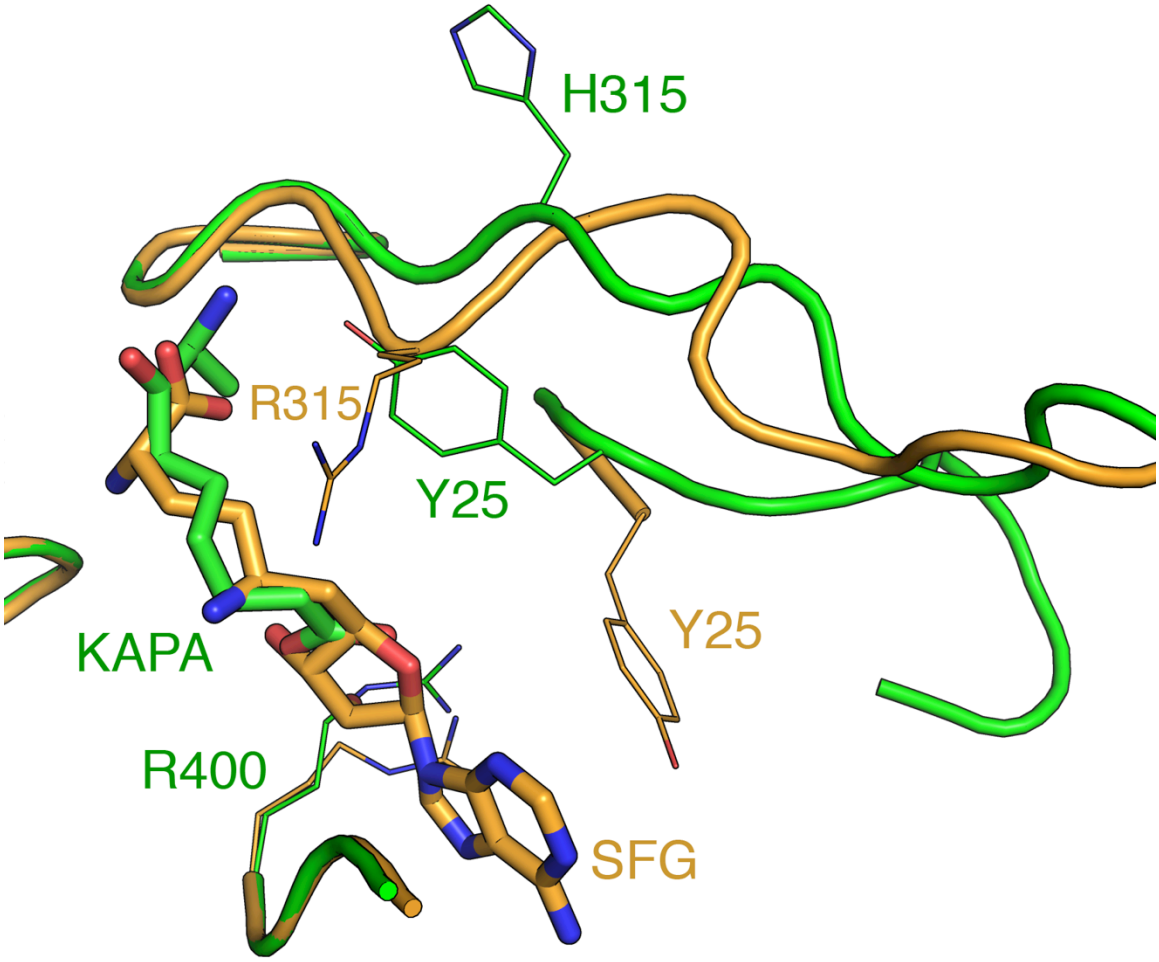
sinefungin in this complex, and this is not indicative of favorable π-stacking. From our own experience, in all of the complex structures we have obtained for BioA (17 structures), Tyr25 is always oriented as seen in the KAPA or *holo* complexes, and never as seen in the sinefungin complex (Figure 2.1). The *B. subtilis* BioA complex (PDB-id 3LV2) is a co-crystal with a H315R mutant, and the histidine occupies the position taken by the side chain of Tyr25 in the wild type enzyme (Figure 2.2). It is possible that the shift in Tyr25 observed in the sinefungin BioA complex is a consequence of this mutation, rather than sinefungin binding.

Figure 2.1. A comparison of the BioA complex with KAPA to other structures.



A) The *holo* (PLP-bound) *Mtb* BioA structure (PDB-is 3TFT)[74]. The enzyme resting state with PLP covalently linked to K283. B) Sinefungin (SFG)-bound structure of *B. subtillis* BioA (PDB-id 3LV2) complex[21]. C) $2F_o$-F_c electron density (1σ) for KAPA, PLP, and Lys283. D) Detail of KAPA binding vicinity.[62]

Figure 2.2 KAPA bound *Mtb* BioA (PDB-id 4CXQ; green) **overlaid with SFG bound H315R *Mtb* BioA** (PDB-id 3LV2; orange, some residues following Y25 are disordered in 3LV2 so they are not drawn in this figure)



2.4. Conclusion

In this chapter, we described the optimized *Mtb* BioA expression, purification, crystallization, as well as the homogeneity validation using DSF. We have prepared *Mtb* BioA at high yield (15 mg/L culture). We obtained protein with nice homogeneity and 100% PLP occupancy after purification. Using the protein we prepared, we were able to increase the crystal diffraction resolution to beyond 1.7 Å. The reproducible production of robust *holo* crystals is necessary to effectively and efficiently validate and characterize bound small molecules that will be identified by fragment screening.

Details of the induced fit mechanism of BioA catalytic cycle are yet to be fully understood; especially regarding how active site conformational changes occur to accommodate two structurally dissimilar substrates. In this chapter, we described the first complex crystal structure of *Mtb* BioA with its substrate KAPA. This structure clearly delineates the substrate binding site, and shows ways in which the protein conformation can adapt in order to accommodate different substrates (Arg 400 inward shift to make polar interactions with KAPA). The *Mtb* BioA-KAPA complex structure has added to an understanding of the enzymatic mechanism of BioA.

Chapter III: BioA lead identification by fragment screening

Acknowledgement: Some material included in this chapter is excerpted from previously published work [Dai, R. et al, 2014. "Inhibition of Mycobacterium tuberculosis Transaminase BioA by Aryl Hydrazines and Hydrazides ". *ChemBioChem*. **15**(4):575-586]. This includes portions of sections 3.2.1, 3.3.2 and figure 3.5. Ran Dai and Barry Finzel contributed equally in writing this published work, with assistance from other authors. Permission to use this copywrited material has been granted by John Wiley & Sons. The fragment DSF screening was done by Todd Geders in the lab of Prof. Barry Finzel. The fragment ITC experiments were done by Ran Dai with aid from Dr. Feng Liu (Prof. Courtney Aldrich lab); the ITC data were processed by Dr. Feng Liu (Prof. Courtney Aldrich lab).

3.1 Introduction

The weakness of ligand-target binding makes the identification of lead fragment compounds challenging.[57] Very sensitive biophysical techniques, such as DSF[82], STD-NMR[83], SPR[65] and X-ray crystallography[84], have been used routinely for screening fragment libraries in FBDD. These biophysical methods have been reviewed in Chapter I. Efforts have been made to combine different biophysical methods[85], and results from different screening methods are rarely alike. It has been recently argued that to better design a fragment library screening protocol, a cascade of biophysical methods should be used.[86]

For our own fragment library screening, a three-stage protocol was designed and used. DSF was first used to identify fragments from the 1000 compound Maybridge Ro3

diversity library that shift the protein melting temperature. Then, macromolecular X-ray crystallography was used as a structure-based validation to ensure specific binding. Finally, multiple biophysical methods (STD-NMR, SPR and ITC) have been used in an effort to determine the binding affinities of each fragment compound to prioritize them for further optimization.

3.2 Experimental procedures for DSF fragment screening

3.2.1 DSF

The DSF screening was performed by Dr. Todd Geders in the lab of Prof. Barry Finzel. In a glass tube, 2475 µL of 2X master mix was made at room temperature by mixing 125.0 µL of 1M HEPES pH 7.5, 50.0 µL of 5M NaCl, 5.0 µL of 5000X Sypro orange stock solution, 2295 µL of water. The 2X master mix was cooled on ice. 310 µL of water was distributed into each well of a single column of a 96-well U-bottom plate at room temperature. 19 µL of water was placed into each well of the white PCR plate held by a room-temperature aluminum block. 1 µL of pure DMSO was placed in each well of the first and last column for DMSO-only controls. Using a multi-channel P10 pipette, 1 µL of each fragment compound was placed into a white PCR plate. BioA protein (25 µL of 13.0 A280) was added to the master mix. 310 µL of the 2500 µL 2X master mix was distributed into each well of a single column of a 96-well U-bottom plate on ice. 20 µL of master mix with protein was added into each well of the white PCR plate and gently mixed. Plates were sealed and run in the Bio-Rad CFX96 using 30 sec dwell time per 1°C temperature increase.

The DSF amplification curve was generated using Microsoft Excel. The DSF first derivative was calculated based on the DSF amplification curve, and the first derivative

curve was also generated using the software Microsoft Excel. The Maximum point of the first derivative was used to determine the T_m shifts.

3.2.2 STD-NMR

STD NMR spectra were obtained at 20 °C using a Bruker 700 MHz NMR spectrometer with a TCI cryoprobe, incorporating Z-axis gradients. Samples contained BioA protein (30 μM) and compound 3E3 (0.1, 0.2, 0.4, 0.6 and 0.8 mM respectively). A one-dimensional pulse sequence incorporating a T1ρ filter was used for the acquisition of STD NMR spectra. The on-resonance frequency was set to 0.8 ppm and the off-resonance frequency was set to 30 ppm. Irradiation was performed using 50-Gaussian pulses with a 1% truncation and a 49-ms duration and separated by a delay of 1 ms to give a total saturation time (T_{sat}) of 1, 2, 3, and 4 seconds for each compound concentration. The duration of the T1ρ filter was 15 ms. STD NMR spectrum was acquired with a total of 6144 transients in addition to 32 scans to allow the sample to come to equilibrium. The spectral width was 8 kHz. Reference spectrum was taken under the same conditions.

The dissociation constant (K_D) was calculated from the binding isotherm of STD initial growth rate using an established model reported by Angulo et al.[87] First, STD amplification factor (STD-AF) was defined as the product of STD ligand intensity and the ligand excess (ε) as shown in equation (1). Then, STD-AF values were plotted with increasing saturation time (t_{sat}) to obtain a curve fit to equation (2) to get STD-AF_{max} and k_{sat} values. Next, an initial STD amplification factor (STD-AF_0) was calculated using equation (3). Finally, STD-AF_0 was plotted against increasing ligand concentration to obtain K_D using equation (4).

$$\text{STD-AF}(t_{sat}) = \varepsilon\,(I_0 - I_{sat})/I_0 \text{ ; where } \varepsilon = [L_0]/[P_0] \quad (1)$$

$$\text{STD-AF}(t_{sat}) = \text{STD-AF}_{max}\,[\,1-\exp(-k_{sat}*t_{sat})] \quad (2)$$

$$STD\text{-}AF_0 = STD\text{-}AF_{max} * k_{sat} \quad (3)$$

$$STD\text{-}AF_0 = (B_{max} * [L])/(K_D+[L]) \quad (4)$$

3.2.3 SPR

All surface plasmon resonance (SPR) experiments were performed on a BIAcore 2000 at 25 ^{o}C. BioA protein substrate was coupled to the flow cells on the CM5 chips by amine coupling following instructions provided in the Amine Coupling Kit from GE Healthcare. BioA was dissolved to make a 0.05 mg/ml solution containing 10 mM Sodium Acetate at pH=5.0 and filtered to remove particulates. The functional groups on the CM5 chip surface were first activated by a 7 minute injection of a freshly prepared 1:1 mixture of 0.4 M 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide and 0.1 M N-hydroxysuccinimide at 10 µl/min; followed by injection of the substrate (BioA) for 7 min at a flow rate of 10 µl/min to > 4000 RU. Remaining activated carboxyl groups on the surface were blocked with 1 M ethanolamine for 7 min at a flow rate of 10 µL/min. A PBS-P buffer containing 1x PBS at pH 7.4 with 0.005% P-20 surfactant and 5% DMSO was used for the following ligand binding assays. In the kinetic binding assay, concentration series in 1:2 dilutions were assessed for binding. A 30-second association phase was followed by a 300-second dissociation phase. All kinetic data were processed using BIAevaluation 3.0 Software. A 1:1(Langmuir) binding model was used to fit the data and calculate the K_D values.

3.2.4 ITC

Isothermal titration calorimetry (ITC) was conducted on a GE MicroCal Auto-ITC 200 microcalorimeter. The titration experiment was performed at 25 °C in ITC buffer (25 mM Hepes [pH 7.5], and 50 mM NaCl). BioA was exchanged into ITC buffer using an Amicon Ultra concentrator, and the final enzyme concentration was determined using the

Bradford assay. In the titration, compounds were individually injected into a solution of the enzyme. Ligand and protein concentrations were optimized. For the optimization of protein concentration, we started with 0.01 μM that had been used to test more potent BioA inhibitors, and gradually increased it. For ligand concentration, a 1:10 to 1:20 ratio of protein to ligand was maintained in the final system concentration. Eventually 0.1 μM BioA and 1.0 μM, 1.5 μM, 2.0 μM fragments were adopted as more concentrated protein gave rise to higher signal to noise in the experiment. The K_A (the association constant in M^{-1}), n (the number of binding sites per monomer) and ΔH (enthalpy) values were determined by ITC. The thermodynamic parameters (ΔG and –TΔS) were calculated using Equation 1:

$$\Delta G=-RT\ln K_A=\Delta H-T\Delta S \qquad (1)$$

Where ΔG, ΔH and ΔS are the changes in free energy, enthalpy, and entropy of binding, respectively, R = 1.98 cal mol^{-1} K^{-1}, and T is the absolute temperature. The affinity of the fragments for BioA is provided as the dissociation constant (K_D =1/K_A).

3.2.5 Crystallography

The crystallization conditions used to produce BioA *holo* crystals were as described in Chapter II. Both soaking and co-crystallization techniques were used to obtain BioA complex structures with fragments. For soaking, BioA *holo* crystals were soaked in reservoir solution containing 5 mM compounds (15% PEG 8000, 100 mM HEPES pH 7.5, 100 mM $MgCl_2$, and 5 mM compound) for 5-60 minutes at 20 °C.

Compound co-crystals were prepared in respective hanging drop vapor diffusion experiments conducted at 20 °C. Protein solution (10 mg/mL in 25 mM HEPES pH 7.5,

50 mM NaCl, 0.1 mM TCEP) was mixed with reservoir solution (9–14% PEG 8000, 100 mM HEPES pH 7.5, 100 mM $MgCl_2$, and 5 mM compound) and a seed solution (a reservoir solution containing crushed BioA crystals) in a 4:3:1 ratio (2 µL BioA protein:1.5 µL reservoir solution: 0.5 µL crushed BioA seed solution). Crystals appear in the drop within 24 hours and grow to their full size in 72 hours. BioA-compound co-crystals were protected by briefly transferring to a cryo solution (15% PEG400, 15% PEG 8000, 100 mM HEPES pH 7.5, 100 mM $MgCl_2$, and 5 mM compound) using an appropriately sized fiber loop of a cryo pin from Hampton Research and then flash frozen in liquid nitrogen. Diffraction data were collected from crystals at 100 K using synchrotron radiation with a Dectris Pilatus 6M Pixel Detector on beamline 17-ID (IMCA-CAT) at APS, Chicago, United States. Data were processed, integrated, and scaled with XDS[76] and SCALA[88] using the autoPROC scripts available at IMCA-CAT. The structures were solved by molecular replacement using Phaser[77] in the CCP4 package[78] using atomic coordinates from PDB code 3TFT as a search model.[25] Refinement and model building was done using REFMAC5[79] and coot[80]. The figures were prepared with PyMOL (The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC.). Structures were superimposed for analysis and display using the shared BioA-PLP overlay method of the DrugSite server.[81]

3.3. Results

3.3.1 Initial screening using DSF

DSF was used for the initial screening of about 1000 compounds from a Maybridge Ro3 diversity library. Compounds that shifted the T_m greater than ± 2 °C were selected for subsequent characterization by macromolecular crystallography. ± 2 °C was selected as the cutoff to afford a 2% hit rate that was desirable for follow-up. Screening allowed

identification of a total number of 21 hits. Nine of these caused upward (stabilizing) T_m shifts while the other twelve caused downward (destabilizing) T_m shifts (Table 3.1).

3.3.2 Binding validation and characterization by X-ray crystallography

An effort was made to obtain complex structures of BioA protein with all 21 fragment-hits identified with DSF with BioA protein. In order to evaluate all 21 compounds rapidly, soaking experiments were performed initially, using compound concentration of 5 mM in reservoir solution with 5-60 minute soak time. By this method, complex structures of fragments 2, 3, 5, 7, and 10 were obtained (Table 3.2). Further increases of the compound concentration to 10 mM did not result in additional complex structures. Later, co-crystallization methods were also applied to all 21 compounds, and one more complex structure (**F9**) was obtained (Table 3.2). Individual crystalline specimens diffracted to varying resolution, but data collection and refinement statistics show that all structures are well-refined and of good quality (Table 3.2). Among the fragment hits that are complexed with BioA in crystal structures, **F2**, **3**, **5**, **7**, **9** are stabilizing hits (they increase BioA T_m) while **F10** is a destabilizing hit (it decreases BioA T_m) (Figure 3.1).

All fragment compounds bind in a hydrophobic pocket adjacent to the PLP cofactor but do not disrupt the internal aldimine that defines the resting state of the enzyme. Electron densities (Figure 3.2) clearly confirm that the covalent bond between Lys283 and the PLP in all complexes remains intact.

The active site of *Mtb* BioA is composed of residues Pro24-Ser34, Ser62-Ala67, Arg156-Asp160, His171-Arg181, Gln224-Gly228, Arg400-Arg403, Met87′-His97′, and Ala307′-Asn322′ (See Chapter II). (Residues marked with a prime are contributed by the other

monomer of the functional dimer). The shape of this active site is as a tunnel towards the inside PLP co-factor. This tunnel is at the interface of the two chains of the BioA homodimer; it is narrow inside but the opening to the surface slightly opens wider. The inner part of this tunnel is the substrate-binding site composed of Ala226, Trp65, Tyr157, Thr318′, Tyr25 and Asp160; the outer section of the tunnel is composed of hydrophobic loops from both chains (His171-Arg181, Ala307′-Met314′, Arg400-Arg403, Met87′-His97′). All six fragments bind in the same portion of this active site occupied by KAPA in the pre-reaction complex (Figure 3.3), although they make different interactions with the surrounding residues. As previously discussed, *Mtb* BioA is a functional homodimer with two active sites. Fragments do not necessarily have the same binding features in both sites, and some fragments are found in only one of the active sites (Table 3.2). There is no obvious difference between the chains that may readily explain this, but it is not unusual to see differences in binding of small molecules in the two "equivalent" positions in BioA crystals,[21] even with respect to the covalent adduct that our group has previously described.[25] More interestingly, the fragments have induced different local conformational changes in BioA. Some hydrophobic residues (aromatic side chains of Trp64, Trp65, Tyr25 and Phe402, for example) interact with fragments in different ways. Below is a structural analysis to highlight some useful information we obtained from each fragment structure.

F2

The co-crystal of **F2** with BioA has a different crystal form ($P2_1$) from the other *Mtb* BioA structures ($P2_12_12_1$); and there are two homodimers in one asymmetric unit (ASU). **F2** binds to only two active sites in one ASU (Chain A and Chain C), and in different orientations in the two sites. As shown in Figure3.4A, they are flipped 180° "head to

tail". In the A chain **F2** interacts with BioA active site mainly through nonspecific hydrophobic contacts; in chain C, **F2** has H-bonds with the amide of Gly-93′ and with the backbone oxygen of Thr318′. The pyridine and thiophene rings of **F2** are prone to favor a rigid co-planer conformation; **F2** cannot form good π-π interactions with both Tyr-25 and Trp64, and so binds so as to maximize the interaction with Trp-64 alone.

F3

F3 occupies both of the two BioA active sites in the asymmetric unit; however, in different binding conformations. When comparing the two binding modes, **F3** is flipped 180 ^{o} from one to the other, but the active site conformations remain the same (Figure 3.4). **F3** is structurally similar to **F2**; the imidazole and benzamide rings in this molecule are co-planar and they both participate in hydrophobic Van der Waals interactions with the aromatic residues of the Tyr25 and Trp64. When we compare the active site of **F3** complex structure to the BioA *holo* structure (pdb code 3TFT), we can see that a 30° shift in Tyr25 is induced to maximize the interaction with the imidazole (Figure 3.4. cyan), and Trp64 is rotated 90° to both make room for the binding and to have a better π-π interaction with the benzamide. The terminal amide group of **F3** interacts with some surrounding water molecules but it does not form any hydrogen bonds directly with the surrounding residues; this fact also enables **F3** to have different conformations in the BioA active site.

F5

The observed binding of **F5** is comparable in both active sites of the asymmetric unit. **F5** is composed of a pyrazole ring, a phenyl ring and a terminal secondary amino group joined by rotationally flexible methylene linkers. The tetrahedral configuration of the

methylene carbon enables a 109° angle between the two planar aromatic systems and allows all elements of the molecule to be involved in specific interactions with the binding site. First, the pyrazole ring makes a π-π interaction with Phe402. The presence of **F5** caused a 0.9 Å inward shift of the Phe402 side chain to better accommodate this interaction. Second, the phenyl ring is in a hydrophobic pocket between Trp64 and Trp65. Trp64 is shifted 1 Å toward the fragment and the sidechain of Trp65 rotates 90^{o} toward the fragment (comparing with the *holo* structure pdb code 3TFT) to make additional hydrophobic interactions. Finally, the secondary amine is 2.9 Å from the PLP phosphate oxygen where it donates a hydrogen bond to the PLP phosphate; it is also 2.9 Å from the Tyr25 phenol oxygen to which it is hydrogen bonded. From a comparison of **F5** to the complex with KAPA, we can see that the geometry of **F5** allows it to occupy much the same binding footprint as KAPA even though these two molecules are structurally very different (Figure 3.3 C).

F7

F7 is observed bound in both active sites in the asymmetric unit. It occupies a sub-site adjacent to the binding site of **F3**. The saturated piperidine can be differentiated from the aromatic benzene ring because the electron density for the piperidine clearly shows the puckered chair conformation. The benzene ring of **F7** resides in a hydrophobic sub-site between Phe402, Tyr407, Trp64 and Trp65. The ketone linking the benzene ring and the piperidine ring serves as a hydrogen bond acceptor paired with the indole NH of Trp65, although the geometry of this interaction is not ideal. The secondary amine in the piperidine is a hydrogen bond donor as it is protonated at neutral pH; it is involved in hydrogen bond interactions with Tyr157, Tyr25 and the PLP phosphate oxygen. No significant conformational changes to the BioA active site occur upon binding **F7**.

F9

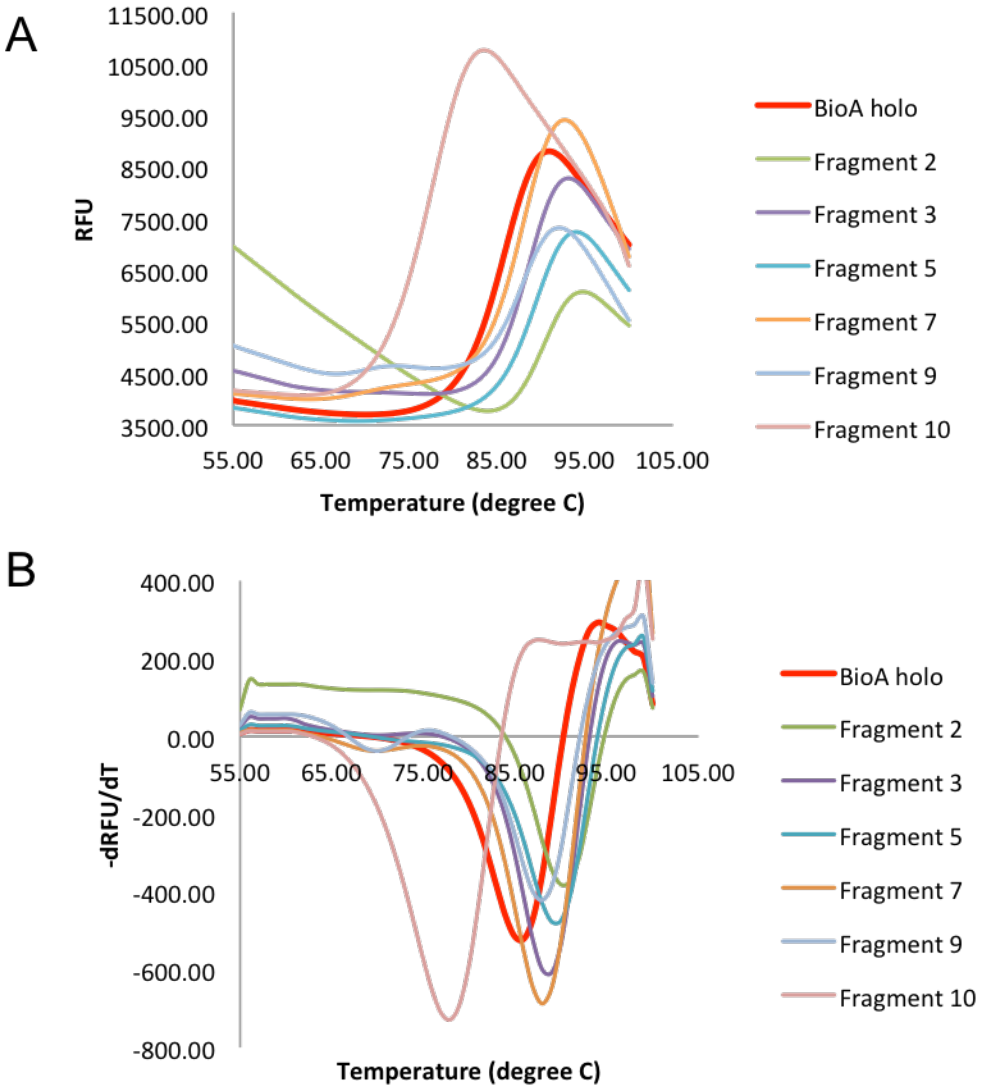
The binding of **F9** was only observed in crystals prepared by co-crystallization, and it is only observed in one active site of the asymmetric unit. The main interaction it makes is a π-π interaction with Trp64. As in complexes with fragments 2 and 3, the Trp64 side chain is rotated approximately 90° to enable this binding conformation and to make this interaction. The lactone is twisted at a 45° angle from the core aromatic ring plane, enabling it to make a stacking interaction with Tyr25.

F10

F10 occupies only one of the two BioA active sites in the asymmetric unit. An unambiguous binding pose can be assigned due to higher density of the sulfur density on the thiazole ring. **F10** binds in direct contact with the PLP, oriented so that the amino group is hydrogen bonded to the phosphate of PLP, Tyr25 OH, Tyr157 OH and Gly316′ O (Figure 3.2 F). This terminal amine acts as a hydrogen donor to the PLP and the Gly316′ O; the third amine hydrogen is directed at the midpoint between the two hydroxyls of Tyr25 and Tyr157 to make a single bifurcated H-bond, much like those that sometimes occur to the midpoint of two carboxylate oxygens. The Tyr25 hydroxyl is also a weak hydrogen bond donor to the nitrogen in the thiazole ring of **F10**. The amino group of KAPA in the complex described in Chapter II makes these same interactions. The binding of **F10** is accommodated with little change to the *holo* enzyme conformational state. The benzothiazole heterocycle lies aside the Lys283 side chain, and between the side chains of Trp64 and Trp65. The two tryptophan indoles are oriented with a 90° angle between them, and the benzothiazole plane almost perfectly bisects that angle, so that comparable hydrophobic contacts are made with each tryptophan. The benzothiazole is just long enough to stretch across the binding pocket and to pack at

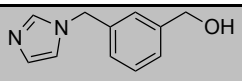
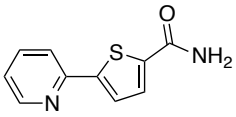
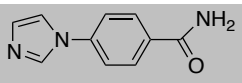
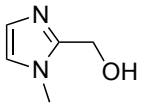
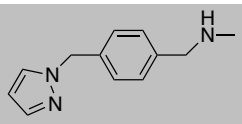
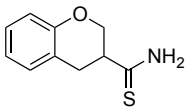
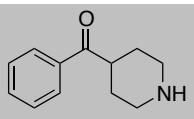
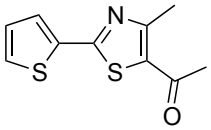
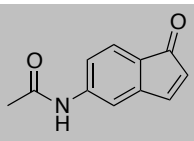
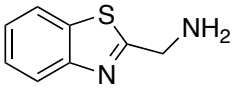
right angles against the face of the aromatic ring of Phe402. Hydrophobic contacts also exist with Met174 and Ala226.

Figure 3.1. DSF results for F2, F3, F5, F7, F9, and F10.



A: DSF melting curve amplification results. B: DSF melting curve derivative results.

Table 3.1. Fragment hits from DSF

Fragment Id	Maybridge Code	Structure	M. W. (Da)	DSF T_m shift (°C)	DSF ΔT_m shift (°C)
F1	CC43209		188.23	90	+5
F2	KM03152		204.25	90	+5
F3	CC18528		187.20	89	+4
F4	CC20809		112.13	89	+4
F5	CC56046		201.27	89	+4
F6	CC24118		193.27	88	+3
F7	SEW04290		189.25	88	+3
F8	SPB02598		223.31	88	+3
F9	TL00757		189.21	88	+3
F10	CC06013		164.23	78	-7

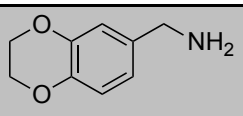
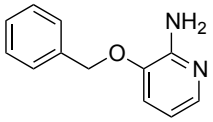
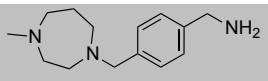
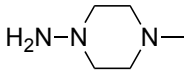
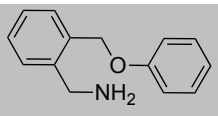
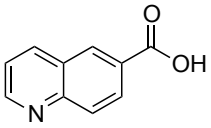
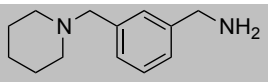
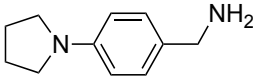
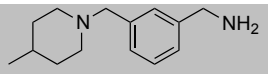
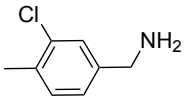
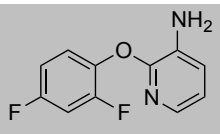
F11	CC01313		165.19	76	-9
F12	BTB08555		200.24	73	-12
F13	CC55813		233.35	73	-12
F14	AC10403		115.18	72	-13
F15	CC58513		213.28	72	-13
F16	CC04501		173.17	71	-14
F17	MO01157		204.31	70	-15
F18	CC30113		176.26	69	-16
F19	MO01158		218.34	68	-17
F20	TL00917		155.62	68	-17
F21	BTB08015		222.19	65	-20

Table 3.2. Crystallographic statistics for complexes with fragments 2, 3, 5, 7, 9 and 10

Data collection	F2	F3	F5	F7	F9	F10
PDB code	4WYA	4WYC	4WYD	4WYE	4WYF	4CXR
Detector	Saturn 944+ CCD	DECTRIS PILATUS 6M	DECTRIS PILATUS 6M	DECTRIS PILATUS 6M	DECTRIS PILATUS 6M	DECTRIS PILATUS 6M
Wavelength (Å)	1.541	1.000	1.000	1.000	1.000	1.000
Space group	$P2_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions						
a, b, c (Å)	62.80, 65.96, 196.79	63.29, 65.98, 204.00	63.10, 66.49, 204.65	62.69, 66.25, 202.33	63.09, 66.09, 203.19	62.94, 66.08, 201.90
α,β,γ (o)	90, 90.19, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Protein chains per ASU	4	2	2	2	2	2
Resolution (Å)	34.39-2.02 (2.09-2.02)	203.99-1.70 (1.76-1.70)	102.32-1.35 (1.40-1.35)	202.33 - 1.75 (1.81-1.75)	101.60-2.25 (2.32-2.25)	100.95-1.70 (1.72-1.70)
R_{merge}	0.146 (0.320)	0.123 (0.460)	0.058 (0.429)	0.066 (0.342)	0.088 (0.171)	0.060 (0.445)
I/σ_I	4.1(1.1)	13.8 (3.2)	18.7 (4.1)	18.4 (4.3)	15.4 (10.3)	17.4 (3.8)
Completeness	65.0% (7.8%)	98.6% (94.4%)	99.0% (99.2%)	99.0% (99.2%)	98.5% (96.7%)	97.6% (96.2%)
Multiplicity	1.9 (1.1)	6.4 (6.5)	6.4 (6.1)	6.0 (6.3)	4.9 (4.3)	4.4 (4.7)
No. observations	135188	609249	1219705	516385	91412	381159
No. unique reflections	70447	94878	189301	85959	38797	86627

Refinement	F2	F3	F5	F7	F9	F10
Resolution (Å)	29.9-2.50	39.7-1.70	32.8-1.35	39.4-1.75	101.6-2.25	100.9-1.70
R_{work} R_{free}	23.3 31.0	17.57 21.34	12.63 15.54	17.58 20.62	19.03 23.56	19.8 22.6
No. atoms	12665	7539	7811	7296	6787	6750
No. water	49	779	937	546	206	232
No. ligand molecule	2	1	2	2	1	1
No. PLP molecule	2	2	2	2	2	2
No. other molecule	0	0	2	0	0	3
Ramachandran plot						
Favored	89.5%	96.6%	97.0%	96.6%	95.2%	96.6%
Allowed	8.6%	2.9%	2.3%	2.7%	2.8%	2.7%
Disallowed	1.9%	0.6%	0.7%	0.8%	2.0%	0.7%
R.m.s deviations from ideality						
Bond lengths (Å)	0.010	0.007	0.006	0.007	0.009	0.008
Bond angles (°)	1.36	1.17	1.14	1.16	1.32	1.32

Figure 3.2. Fragments in BioA active site with 2Fo-Fc (1σ) electron density display.

A: **F2**, B: **F3**, C: **F5**, D: **F7**, E: **F9**, F: **F10**

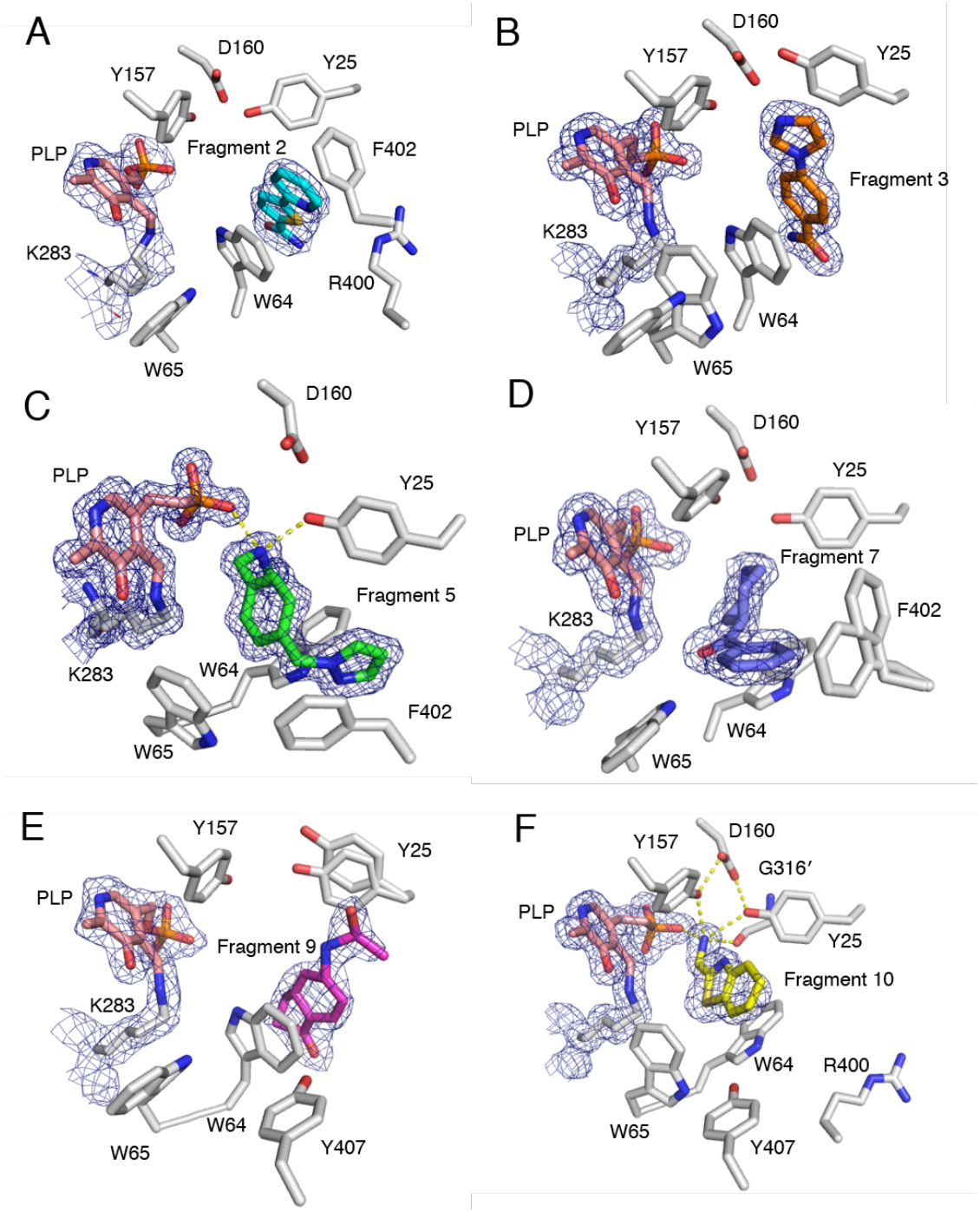


Figure 3.3. Fragment structures overlaid on KAPA (shown in grey) to emphasize sidechain conformational differences induced by different ligands. A: **F2**; B: **F3**; C: **F5**; D: **F7**; E: **F9**; F: **F10**.

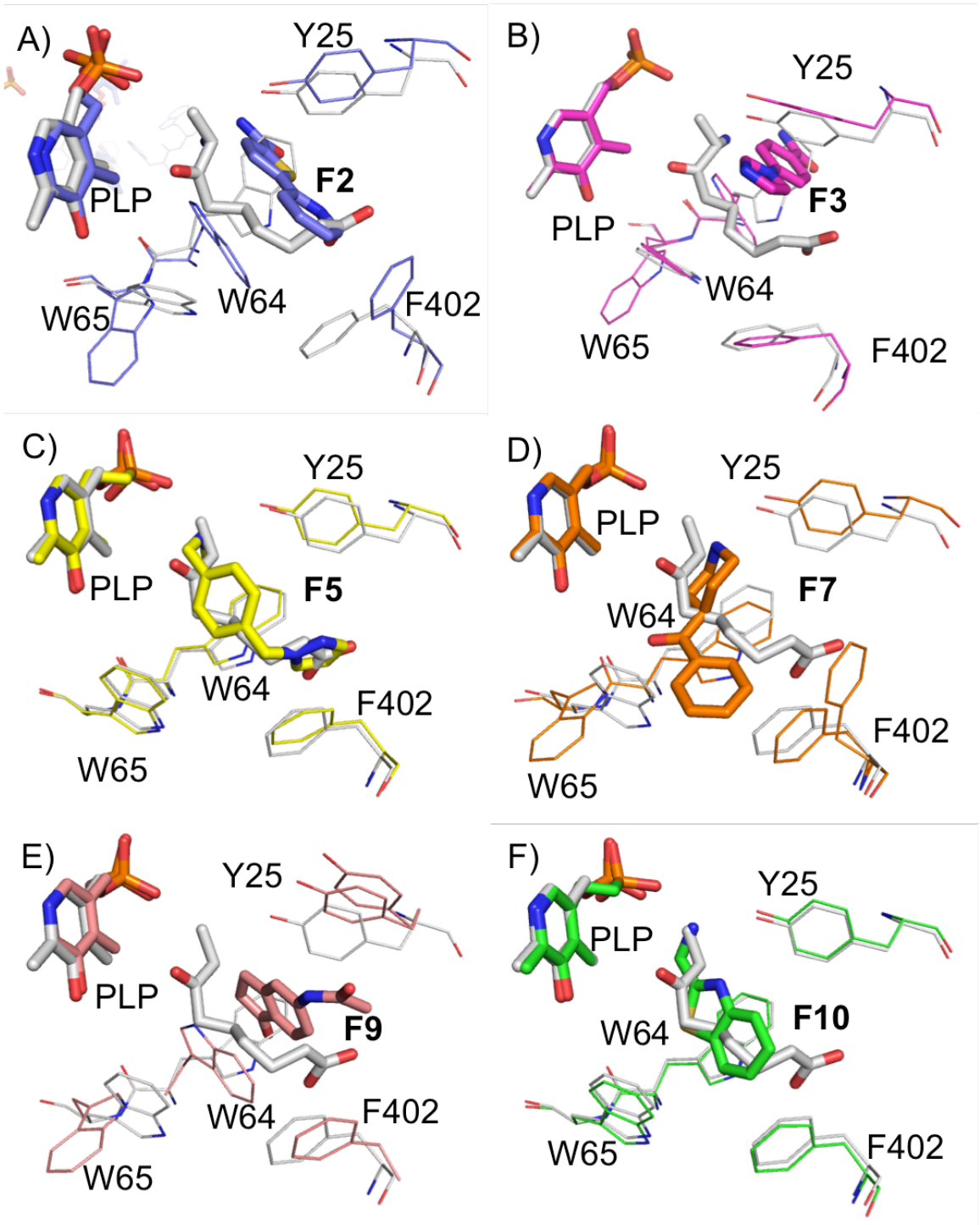
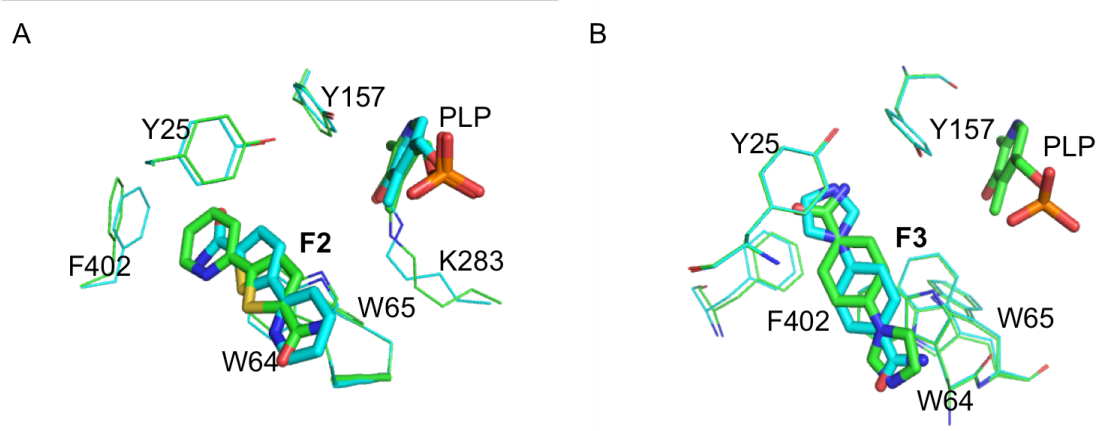


Figure 3.4. (A) F2 binding conformations in BioA active site. The binding conformation in chain A is shown in green and that in chain C is shown in cyan. **(B) F3 binding conformations in BioA active site.** The binding conformation in chain A is shown in green and that in chain B is shown in cyan.



3.3.3 Assessment of fragment binding affinities

DSF is a qualitative screening tool for hit identification; it provides evidence of inhibitor binding, but cannot provide quantitative assessments of ligand binding affinity. Likewise, X-ray crystallography provides detail regarding ligand binding; however, it has been shown that some weak binders (K_D of 10^{-2}M) can be characterized in a complex X-ray crystal structure.[89] One piece of important information we cannot obtain from DSF or X-ray crystallography is the fragment binding affinity. Fragments usually have comparatively low binding affinities to protein targets (as we discussed in Chapter 1); however, choosing a fragment with comparatively strong and specific binding as a starting point is crucial for accelerating future inhibitor design. For this reason, several different biophysical methods were explored for the purpose of determining the binding affinities of the fragments.

STD-NMR was first used in attempt to determine the binding affinities of the fragments. As data collection for this experiment consumes a large amount of protein and compound samples, we only chose **F10** to develop experience with this method. **F10** has three pairs of ^{1}HNMR signal with chemical shifts at 4.54ppm, 7.43ppm and 7.95ppm respectively (Figure 3.5 a). We can identify their signal from the STD-NMR experiment (Figure 3.5 c). Proton signals in all of these 3 groups were used to calculate a K_D value, but K_D values estimated using different proton signals do not agree with each other. The K_D was calculated as 17.2 μM from ^{1}HNMR signal with chemical shifts at 4.54 ppm, 4.56 mM from ^{1}HNMR signal with chemical shifts at 7.43 ppm, and 58.8 mM from ^{1}HNMR signal with chemical shifts at 7.95ppm. We turned to other methods to quantify the K_D. The possible reason for the disagreement is further addressed in the discussion section.

SPR was our second option to estimate K_D values. BioA was immobilized on a CM5 Biacore chip using procedures described in section 3.2.3. Different concentrations of fragment compounds (31.25μM, 62.5μM, 125μM, 500μM, 1mM and 2mM) were flowed through the BioA surface channel and relative responses (RR) were collected for each experiment. A 1:1(Langmuir) binding model was used to fit the data and calculate the K_D values. The compounds were not well-behaved in the SPR experiments and the K_D values estimated from repeated experiments did not agree well.

Isothermal titration calorimetry (ITC) was not chosen to be our primary fragment screening tool because it requires relatively large amounts of protein and ligand for each experiment. A big advantage of ITC is that it provides the thermodynamic parameters of protein-fragment interactions in solution. We have optimized the experimental conditions and realized that for the fragment hits, a relatively high protein concentration (0.1 mM) is necessary to achieve a high signal-to-noise ratio. Because of the high concentration (0.1mM) of BioA required in ITC experiments with low affinity fragment compounds, ITC experiments were only completed for the six fragment compounds confirmed to bind crystallographically. Three different fragment concentrations were used (1 mM, 1.5 mM, and 2 mM) (Figure 3.6), and all experiments were completed in duplicate with good reproducibility. The data with the best fitting curves were used to generate the thermodynamic binding parameters ΔH, ΔG, -TΔS, and K_D summarized in Table 3.3.

The ITC results showed that all six fragments have single or double-digit micromolar binding affinities, which is a good starting point for drug discovery. Taking into account the fact that the fragments are relatively small molecules, ligand efficiency (LE) was calculated for each fragment. All of fragments have good ligand efficiency (LE > 0.4). In particular, **F10** has ligand efficiency as high as 0.59, showing high potential for

optimization into lead compounds. Calculated solvent partition coefficients often parallel small molecule binding in hydrophobic pockets, so calculated logarithm partition coefficients (cLogP) are included in Table 3.3.

Figure 3.5. STD-NMR spectra for F10. (a) is the reference 1H spectrum; (b) is the STD-NMR 1H spectrum; (c) is the difference spectrum which shows the ligand signal which confirmed that **F10** binds.

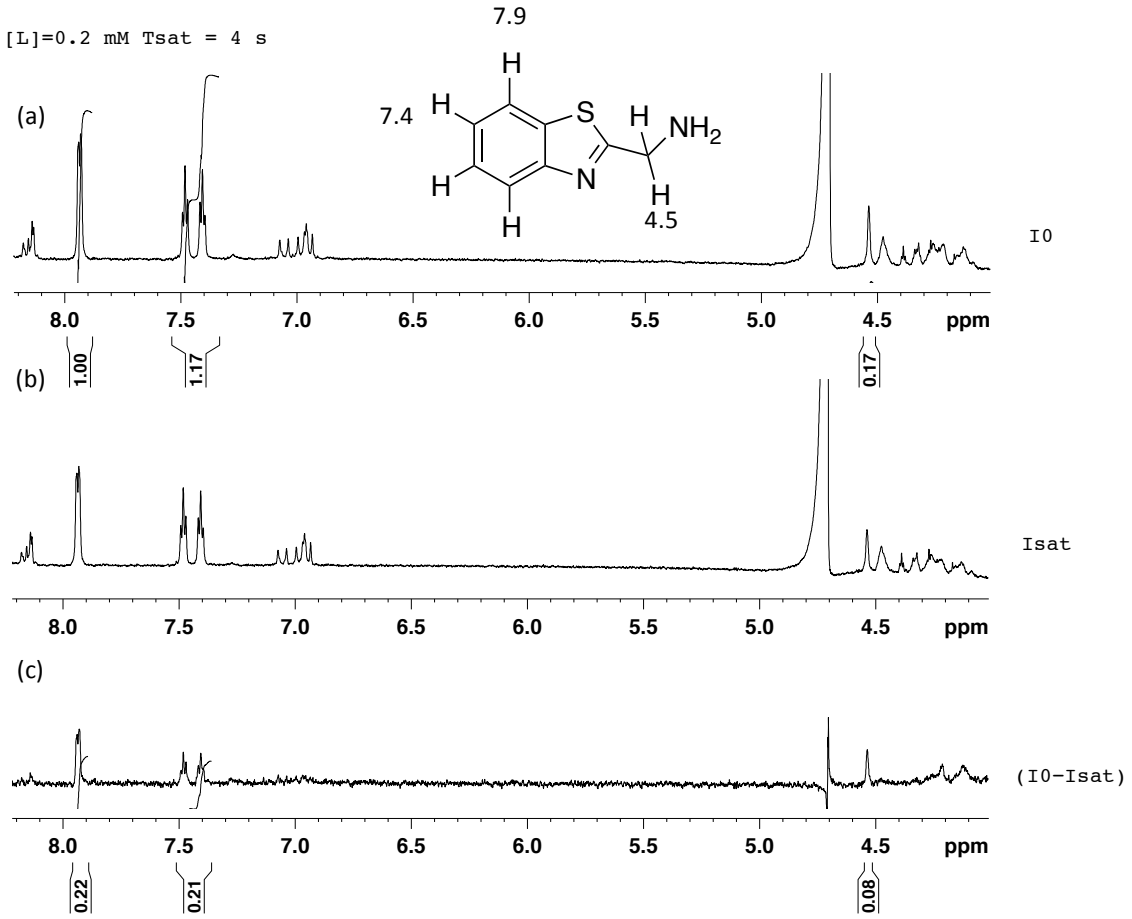
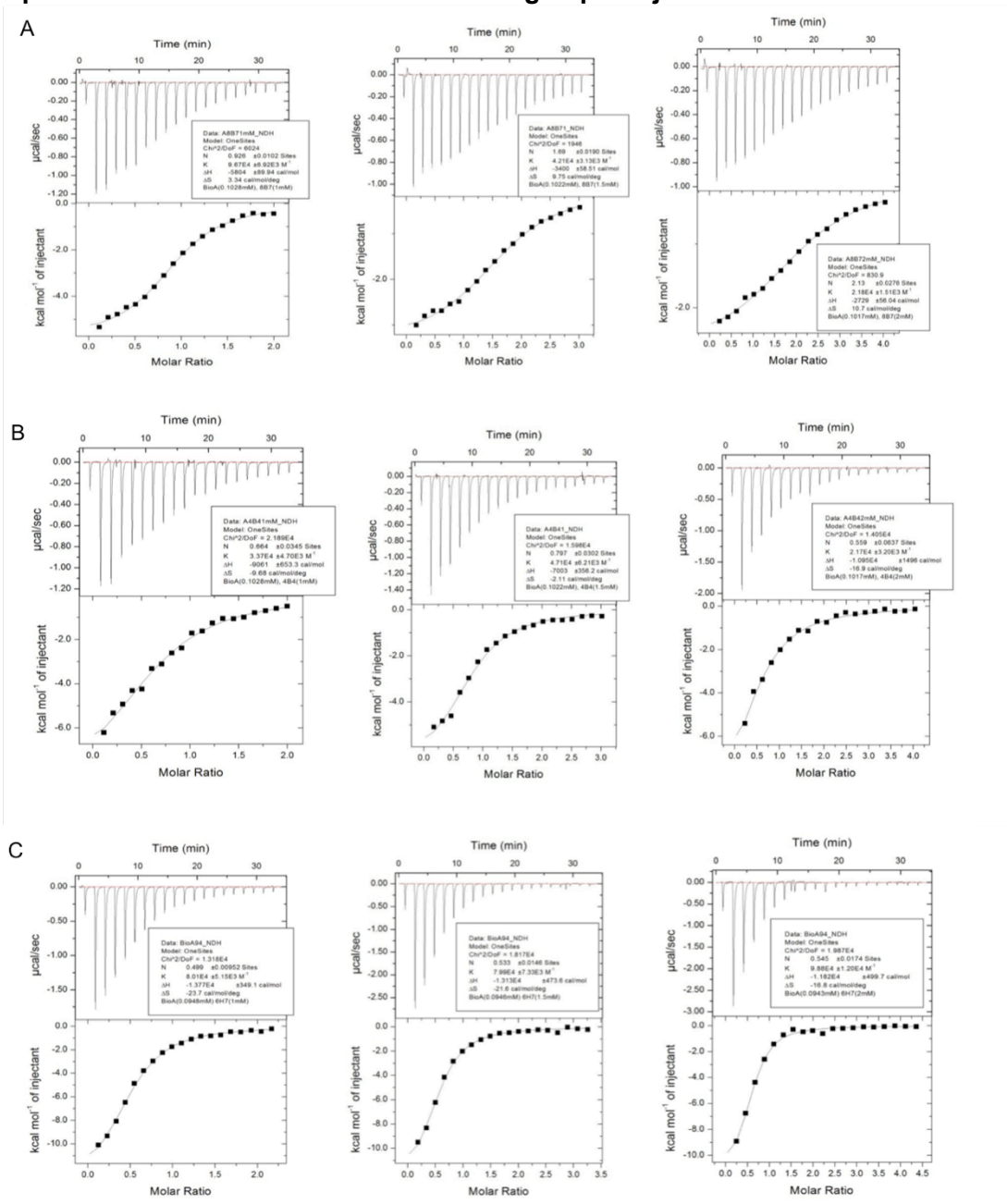
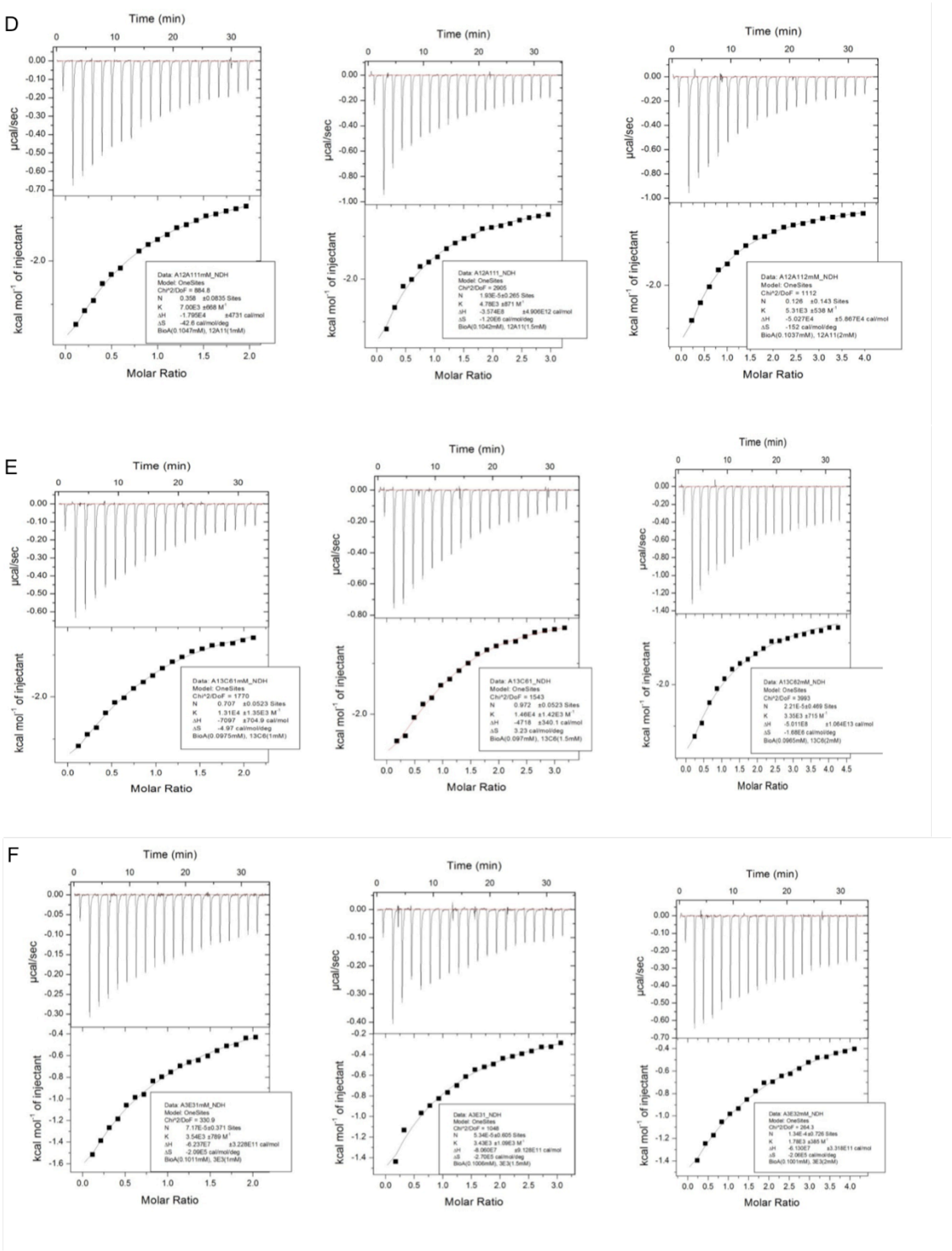


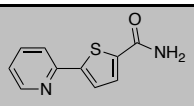
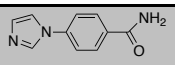
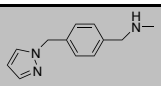
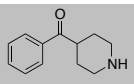
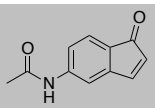
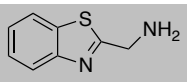
Figure 3.6. ITC results for fragments 2 (A), 3 (B), 5 (C), 7 (D), 9 (E), and 10 (F) represented as a function of heat exchanged per injection.





The upper figures show the time dependence of the electric power (µcal/sec) to maintain constant temperature of the sample after each injection. The lower figures show the heat per mole of injectant vs. the molar ratio of the ligands and protein in the system. Figures from left to right are for each compound using 1mM, 1.5mM and 2M compounds.

Table 3.3. Thermodynamic parameters of fragment binding with BioA

	F2	F3	F5
Structure			
ΔH, kcal mol^{-1}	-5.4±0.08	-6.1±0.15	-13.5±0.30
K_A, 10^4M^{-1}	13.1±1.0	6.9±0.5	8.5±0.5
ΔG, kcal mol^{-1}	-6.9	-6.6	-6.7
-TΔS, kcal mol^{-1}	-1.5	-0.5	6.8
K_D, (μM)	7.6	14.6	11.8
cLogP	0.7	1.3	1.6
No. H bonds with BioA	0	1	3
Ligand efficiency*	0.53	0.47	0.45
	F7	**F9**	**F10**
Structure			
ΔH, kcal mol-1	-4.2±0.35	-3.1±0.16	-1.3±0.07
KA, 104M-1	2.3±0.3	4.6±0.6	5.9±0.7
ΔG, kcal mol-1	-5.9	-6.3	-6.5
-TΔS, kcal mol-1	-1.7	-3.2	-5.2
KD, (μM)	43.4	21.6	16.9
cLogP	1.2	1.2	1.0
No. H bonds with BioA	2	1	4
Ligand efficiency*	0.42	0.45	0.59

*Ligand efficiency was calculated using the equation LE = (ΔG)/N where N is the number of non-hydrogen atoms in the molecule.

3.4 Discussion

3.4.1 Difficulties in fragment binding affinity determinations

Fragments are often weak binders; the binding affinity of fragments, especially in the form of ligand efficiency, is a very important piece of information that can help guide molecular designs based on fragments. STD-NMR, SPR and ITC are all biophysical techniques that can theoretically determine the binding affinities of low affinity ligands. However, in any particular case, any one of these techniques may prove unreliable. In the case of BioA, both STD-NMR and SPR experiments failed to provide valid and reproducible K_D values. In STD-NMR, the saturation transfer difference signal is not the same for protons with different chemical shifts. The distance of the fragment-associated proton to the protein is another factor: the closer the proton is to the protein, the stronger a STD signal will be. This feature can provide useful information regarding ligand binding. However, it has been argued that protons with larger STD-AFs will give rise to a higher apparent K_D [87]. In the case of BioA, using different proton signals results in different apparent K_D values.

SPR been used extensively to study fragment binding[65–67,89]. In the initial design of this experiment, we first used common amine coupling to immobilize BioA on the reflective chip. It is likely, however, that reagents used in the amine coupling interfere with the PLP cofactor binding in the BioA active site. To be more specific, the ethanolamine used to block the remaining exposed active carboxylates on the surface has potential to react with the aldehyde of the PLP, releasing it from the protein active site. Similarly, some fragment or inhibitor analogs may contain chemical groups that potentially react with PLP. Irreversible modification of the protein or the surface to which it binds will impair data quality, since the quality of the protein surface cannot be restored between multiple

runs. For these reasons, we decided that SPR was not a suitable assay for determining our fragment binding affinities in this case.

Isothermal titration calorimetry (ITC) is a biophysical technique often used to obtain thermodynamic parameters of ligand binding. It can be sensitive within the range required for analyzing fragment binding, and, when experiments are carefully conducted, it can produce robust outcomes. We have tested all of our fragments in replicate experiments with different ligand-protein ratios. K_D values estimated using different conditions agree quite well, but there tends to be broader distributions of enthalpic energy (ΔH) and entropic energy (-TΔS) in the total energy when using lower ligand:protein ratios. This is due to the relative weakness of the fragment binding, as well as low molecular weight of the fragment ligands, which resulted in relatively low signal to noise ratios. Still, from the ITC results, we can quantitatively analyze if the fragment binding is enthalpy or entropy driven with reasonable reliability. This piece of information may be very important for later inhibitor design.

3.4.2 Thermodynamics parameters of fragment binding

Both enthalpic and entropic energies contribute to protein-ligand binding. Polar interactions with the “hot spots” (structurally conserved spots that are enriched in polar residues and can form strong hydrogen bonds with multiple ligands)[90] are enthalpic contributions. Specific stabilizing non-polar interactions, such as those arising from π-π or π-proton interactions are also enthalpic contributions. Solvation of hydrophobic ligands and apolar binding pockets account for most entropic contributions. Most of these fragments have similar calculated logPs near 1.0, which suggests that little change in the entropy of solvation occurs when fragments bind. A primary difference in the TΔS component of binding may arise in cases where conformational states achievable by

rotatable bonds in solution are excluded from binding to the BioA binding site. Such restrictions would be reflected in an entropic penalty evident in thermodynamic data.

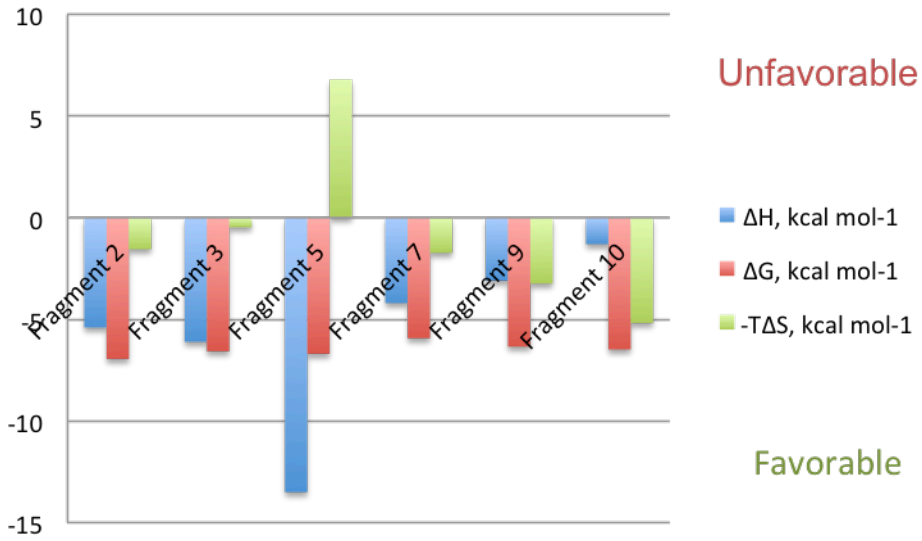
Because fragments are small molecules (less than 22 heavy atoms), and few rotatable bonds, merely entropic energy of solvation (1kJ/mol per heavy atom) is not sufficient to drive binding. Enthalpy driven energies, such as hydrogen bonding interactions with hot spots (over 5 kJ/mol per H bond) and π-π interactions (1 kJ/mol per atom) contribute significantly to good fragment binding hits. The enthalpy favored binding relies largely on specific hydrogen bond interactions with the hot spots, that are unique features of the binding site. Thus, molecules with enthalpy driven binding are in general superior to entropy driven hits as "good starting points" in drug discovery[91].

ITC analysis enables us to quantify enthalpic (ΔH) and entropic (-TΔS) contributions to the total binding energy (Figure 3.7). From the thermodynamic parameters, most of the fragments have moderately favorable entropic interactions with BioA active site (-TΔS around -3, Figure 3.7). The LogPs calculated for all fragments are similar and slightly greater than 1, from which we may surmise that they have similar entropies of solvation and a slight preference for a hydrophobic binding site. **F5**, however, is unique in the large unfavorable entropy upon binding to BioA (-TΔS 6.8, Figure 3.7). Compared to **F5**, fragments **2**, **3**, **7**, **9** and **10** are more rigid compounds without rotatable linkers; their binding results in little net entropy change, **F5** has more rotatable bonds and it can be more flexible in a solvent environment than in the binding pocket; the conformational constraints imposed upon binding result in a fairly large entropic penalty. This disfavored entropy is more than offset, however, by very favorable enthalpic energy of binding by this fragment. **F5** makes 3 strong hydrogen bonds with a hot spot composed of the PLP O, Tyr25 O and Gly316′O. **F7** and **F10** also have more than 2 hydrogen bonds with

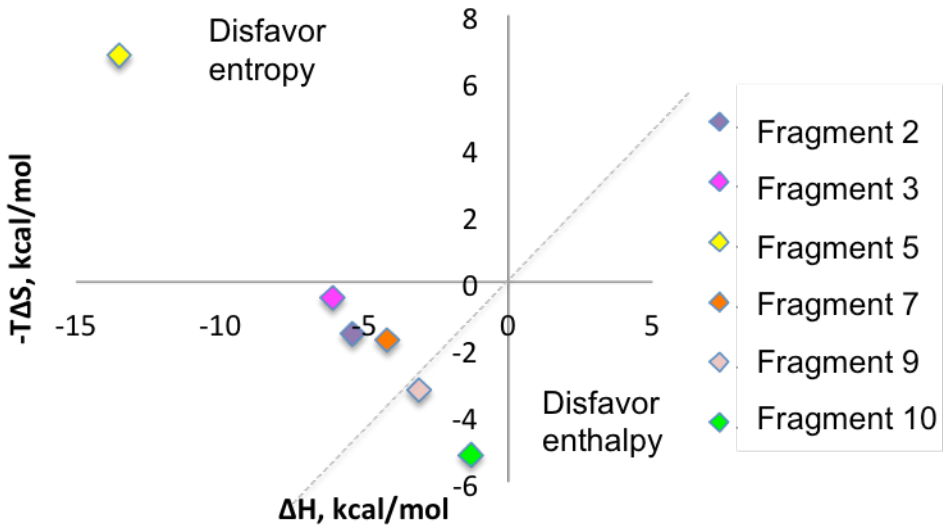
residues in BioA active site, but because the polar interactions are not in optimum geometries, enthalpic interactions with BioA appear to be weaker (Table 3.3). From our analysis of the ITC-derived thermodynamic parameters, we can characterize **F5** as an enthalpy driven fragment hit, while the other fragments are enthalpy-entropy driven (Figure 3.7). Given the recommendation that preference be given to enthalpy driven binding[91] **F5** appears to be a better starting point than the other fragments for further molecular design.

Figure 3.7. Histogram of thermodynamic parameters of the fragment bindings

A



B



A: Histogram of the ΔH (blue), ΔG(red) and -TΔS(green). Negative values shows energy favorable for binding. B: Plot analyzing the enthalpic and entropic component of the binding energy. The binding of **F5** (yellow) is enthalpy driven, and the other fragments are enthalpy-entropy driven.

3.4.3 BioA active site conformational changes upon fragment binding

These structural results reveal that the BioA active site can adopt different conformations to accommodate binding and improve interactions with particular small molecules. In other words, upon binding with different non-covalent fragment ligands, different BioA binding site conformations were induced. How these conformational changes induced from the *holo* BioA state correlate with inhibition becomes an interesting question. In the study of the BioA complex structure with KAPA bound in its pre-catalytic site (described in detail in Chapter II), we observed no reorganization of the key residues that interact with fragments (Y25, W64, W65). These residues are believed to remain in the same conformation throughout the catalytic cycle. Key active site side chains occupy different combinations of conformations in different fragment complexes, as summarized in Table 3.4 and Figure 3.8. Interestingly, in six fragment complexes, the same configuration of these few sidechains is only observed twice. **F10** binds and induces no changes from the *holo* structure. Fragments 3 and 9 induce the same altered state. Some conformational attribute of each of the other complexes is unique.

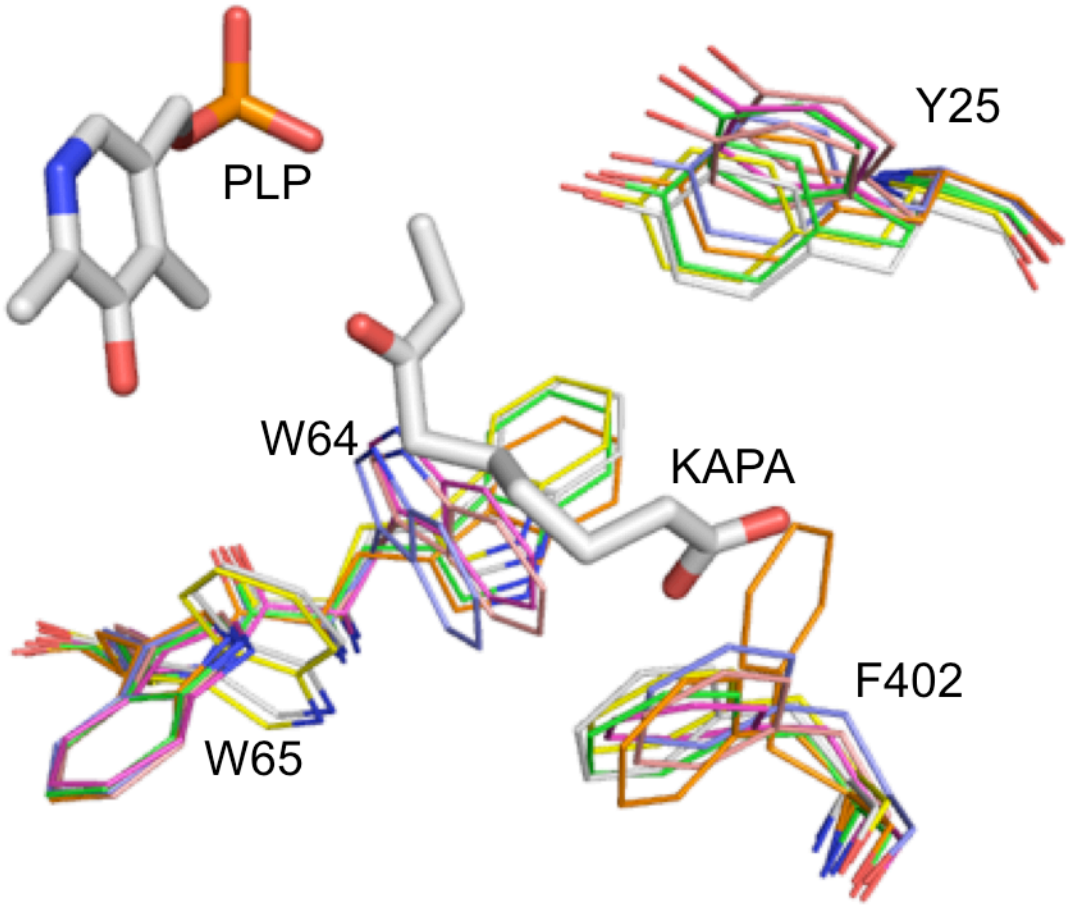
These versatile conformational adaptations would make BioA a challenging target for design of inhibitors based solely on the *holo* crystal structure. Computer-assisted methods such as molecular mechanics (MM) or molecular dynamics (MD), largely rely on assumptions that the protein conformation is unchanged upon ligand binding. The fragment bound BioA structures sampled as a part of this study represent many possible conformations of the BioA active site. Each should be considered as a possible bound state in future computational docking, scoring, or virtual screening. It will be important to match the ligand to the appropriate protein conformation, so that all specific interactions may be modeled accurately.

Table 3.4. Conformation clustering for key residues in the fragment-binding site.

	Y25	W64	W65	F402
BioA holo	0 and +30	0	0	0
F2	0	-90	0	0
F3	+30	-90	0	0
F5	0	0	+90	0
F7	0	0	0	0 and +45
F9	+30	-90	0	0
F10	0 and +30	0	0	0

Side chain conformation of holo BioA (pdb-id 3TFT) was defined as position 0. Conformational changes of residues in each fragment complex structure is defined by its shifting angle and shifting direction (clockwise '+' or counterclockwise '-').

Figure 3.8. Side chain conformations of key residues in BioA fragment binding sites showing in different colors.



Cyan: BioA holo; Purple: **F2** complex; Magenta: **F3** complex; Yellow: **F5** complex; Orange: **F7** complex; Pink: **F9** complex; Green: **F10** complex. KAPA (grey) is included as a ligand binding position reference.

3.5 Conclusion

In this chapter, we described a three stage protocol for fragment screening. From the initial DSF screen, 21 out of 1000 compounds were selected (hit rate 2.1%). Further validation using X-ray crystallography confirmed that 6 out of the 21 DSF hits bind in the active site of BioA (hit rate 0.6%). No other binding sites for fragments were found. Efforts have been made using multiple biophysical methods to obtain binding affinities of the 6 fragment hits and ITC provided such information. The binding of each fragment has been characterized structurally and specific molecular features have been identified that help to explain the origin of binding affinity for BioA. This survey of multiple ligand bound complexes also has exposed considerable flexibility in the BioA active site that should be of use in interpreting future affinity data as the design of more potent and selective BioA inhibitors progresses.

Chapter IV: Structure-based Optimization of Fragment Binding

Acknowledgement: Some material included in this chapter is excerpted from previously published work [Dai, R. et al, 2014. "Inhibition of Mycobacterium tuberculosis Transaminase BioA by Aryl Hydrazines and Hydrazides ". *ChemBioChem*. **15**(4):575-586]. This includes portions of sections 4.2.5, 4.2.7, 4.2.8, 4.3.4, 4.4.1, 4.4.2 and figures 4.6, 4.7 and 4.8. Ran Dai and Prof. Barry Finzel contributed equally in writing this published work, with assistance from other authors. Prof Courtney Aldrich contributed to part of the writing of section 4.4.1 and 4.4.2. Permission to use this copywrited material has been granted by John Wiley & Sons. The kinetic assays of **F10-1**, **F10-2** and **F10-3** were done by Daniel Wilson from the lab of Prof. Courtney Aldrich. The ITC experiment of **F5-1** was done by Dr. Feng Liu.

.

4.1 Introduction

After the initial fragment screening, there are several options for lead optimization that we have discussed extensively in Chapter I. Fragment extension (or fragment evolution) is ideally applicable when a fragment shown to bind in one site of the target can be evolved by extension to make additional favorable interactions in an adjacent sub-site for more favorable binding (Figure 4.1) [35]. Using this strategy, new chemical entities are designed based on the structural information of fragment binding to exploit additional or improved interactions with the existing binding site. There are many examples involving a variety of targets where a fragment extension strategy has been employed. [31,91-92]

Traditional SAR studies have been widely used in conjugation with fragment based approaches,[93-94] but with the detailed fragment binding information from X-ray

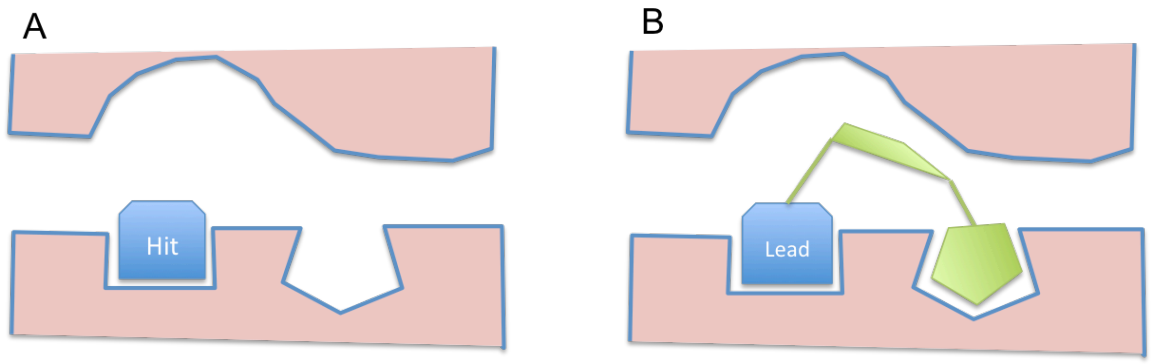
crystallography, molecular designs can be much more rational. From the *Mtb* BioA structures we can see that all the fragments occupy portions of the active site near the PLP cofactor. This active site is a deep, tunnel-like pocket composed of loops contributed by both molecules of the homodimer including Pro24-Ser34, Ser62-Ala67, Arg156-Asp160, His171-Arg181, Gln224-Gly228, Arg400-Arg403, Met87′-His97′, and Ala307′-Asn322′ (Figure 4.2) [21]. None of the fragments fill this site completely, so there are plenty of sub-sites adjacent to the fragments for extension. The high-resolution fragment complex X-ray structures reveal details of the BioA side chain conformations, and have allowed us to identify specific protein-ligand interactions that can be improved.

The Maybridge Ro3 core library we used for fragment screening was composed of a relatively small number of molecules meant to be representative of a diverse chemical space; many analogs of these molecules are also commercially available. Often, even very small changes to fragment molecules can result in large improvements in potency and ligand efficiency[33,95], so the purchase and characterization of available fragment analogs was undertaken as an efficient and economical method for initial fragment optimization, and to explore Structure Activity Relationships. Such “SAR by commerce” can provide a means to modify fragment hits at different locations with very tentative changes, but without expensive chemical synthesis.

In this chapter, we report our discovery of *Mtb* BioA inhibitors a using structure-based fragment extension strategy. First the fragment bound structures were studied to identify opportunities for fragment extension consistent with binding position. Second, SAR by commerce was conducted to identify molecules with higher binding affinity and to explore SAR. In selecting compounds for purchase, both a “rule of 5” filter and structure-

based filters were applied. Finally, the binding of selected compounds was evaluated by DSF and, when co-crystals could be prepared, X-Ray crystallography. Confirmed lead compounds were further characterized by biochemical activity or ITC.

Figure 4.1. Fragment optimization by fragment extension.



(A) Fragment hit as bound in a sub-site; (B) lead compound designed to extend into a neighboring sub-site (fragment extension). Figure adapted from Rees *et al*. [35]

4.2. Experimental Procedures

4.2.1. Compound sources

Compounds used in this chapter are purchased from Chembridge online chemical store (Hit2Lead.com) or requested from National Cancer Institute Developmental Therapeutics Program (NCI DTP). All chemicals used meet the high quality standard of 100% NMR identification. Compounds from Hit2Lead.com were at least 95% pure as evaluated by LC-MS and the compounds from NCI DTP were at least 90% pure by LC-MS. Compound IDs provided in Table 4.1 and 4.2 are Chembridge catalog IDs or NCI DTP compound IDs.

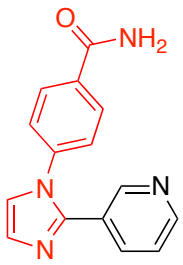
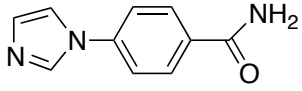
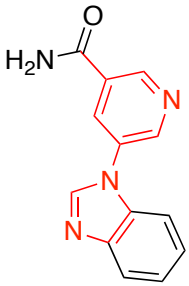
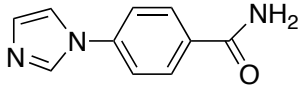
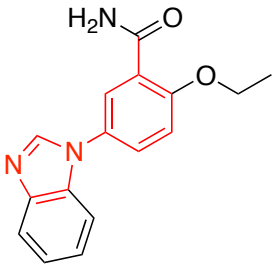
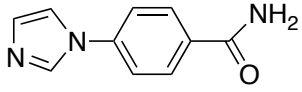
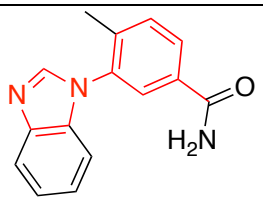
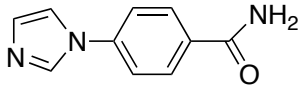
4.2.2. Selecting analogs of fragment structures

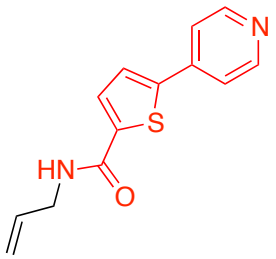
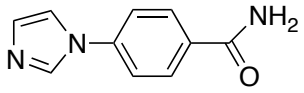
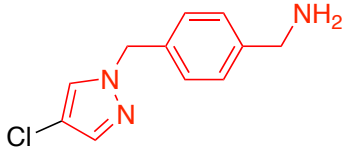
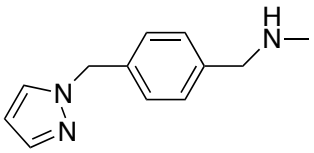
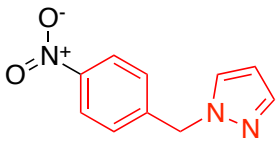
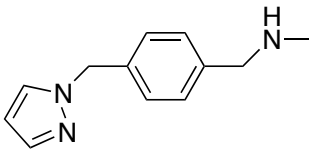
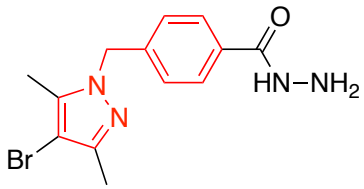
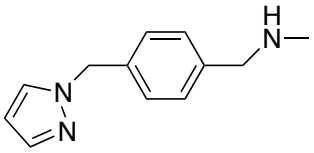
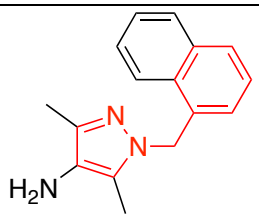
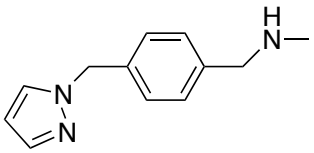
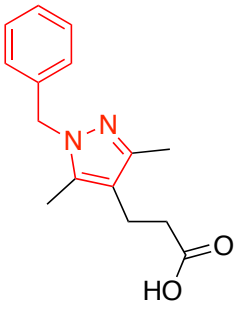
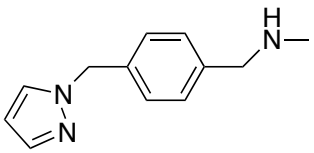
A substructure search method was used to identify commercially available analogs that contain core structural features of each fragment hit. The fragment core structures were selected based on the X-ray complex structures. They maintain the shape, hydrophobicity and polarity features of the original fragments; they are assumed to bind in the BioA active site in the same conformation as the original fragment hits. Scifinder and PubChem database servers were used to search for candidate compounds. Substructure searching was used to identify fragment analogs from extension. The core structures (shown in red in Table 4.1) were used to search molecules containing each core. In addition, a similarity search was used to identify other simple modifications that might potentially increase binding affinity. Only commercially available compounds were selected from the search results to build the initial candidate library for further screening.

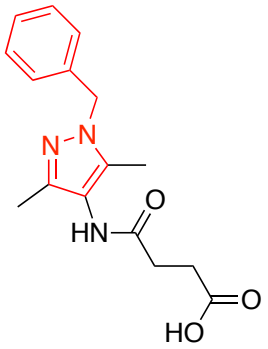
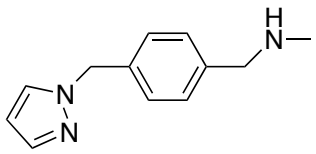
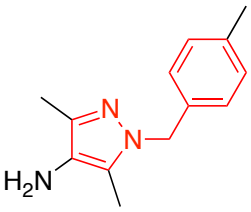
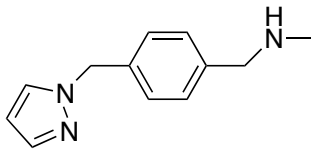
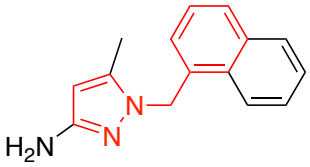
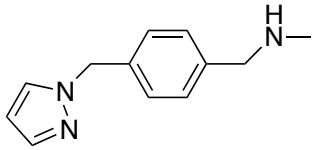
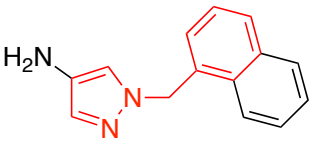
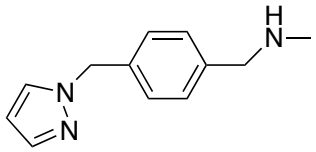
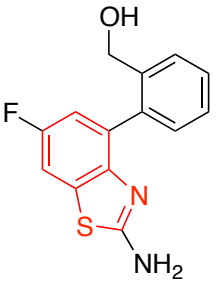
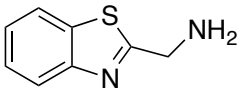
4.2.3. Filtering prospective analogs

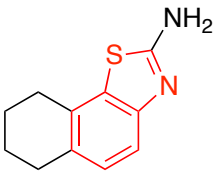
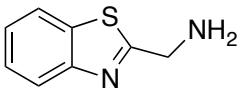
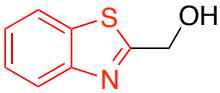
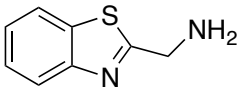
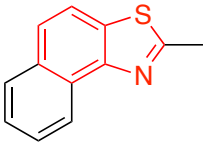
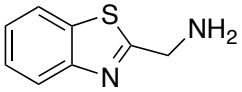
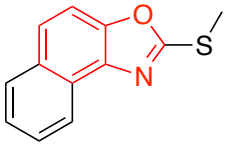
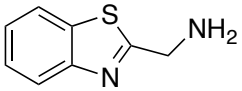
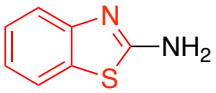
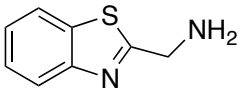
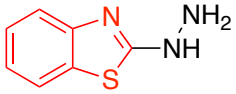
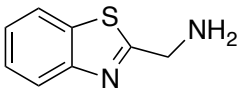
50 compounds (Table 4.1) selected from Scifinder and PubChem databases were screened using the Lipinski's Rule of 5 filter in the Schrödinger Suite. To do this, structures of the candidate molecules were input in Schrödinger software, a Ligprep was performed with Lipinski's rule selected as the pre-filter. All the 50 small compounds passed through the Lipinski's rule filter. Then 3D compound structures were built using the Prodrug Server[96], and each compound was superimposed onto the fragment in the BioA binding site to align the conserved substructure to that in the X-ray structure so that molecules that introduce obvious steric conflicts could be eliminated.

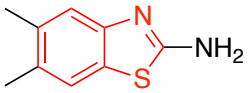
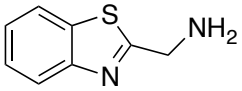
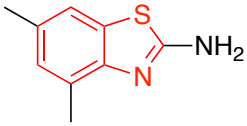
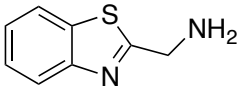
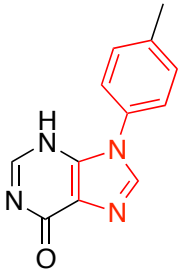
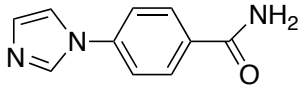
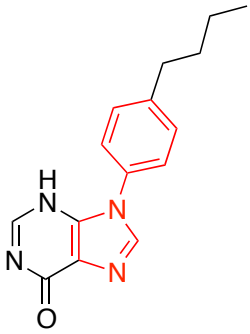
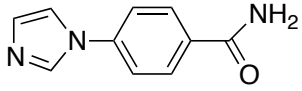
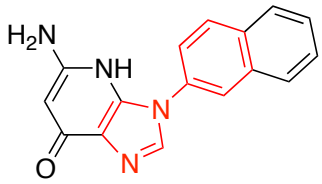
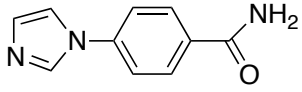
Table 4.1. Compounds selected for initial SAR by commerce. The fragment core in each compound is colored red.

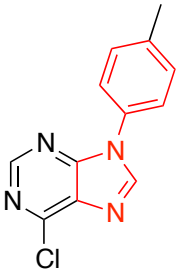
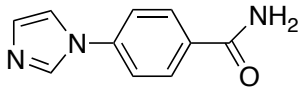
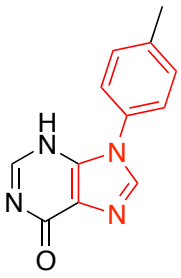
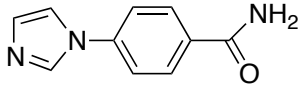
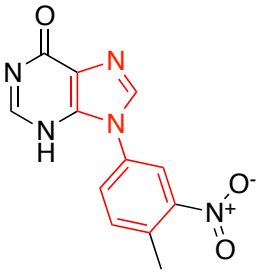
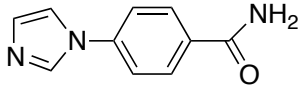
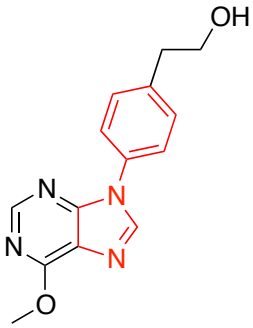
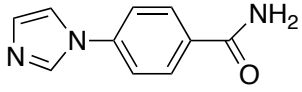
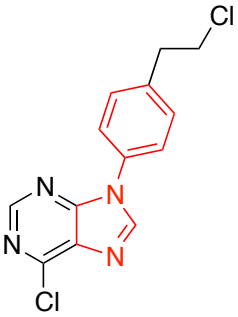
Compound ID	Structure	Template	Descriptions
10896690		**F3** 	M.W. : 264 H donor: 1 H acceptor: 3 cLogP: 1.24
13846515		**F3** 	M.W. 238 H donor: 1 H acceptor: 3 cLogP: 1.22
94450964		**F3** 	M.W. 281 H donor: 1 H acceptor: 3 cLogP: 2.70
81034737		**F3** 	M.W. 251 H donor: 1 H acceptor: 2 cLogP: 2.66

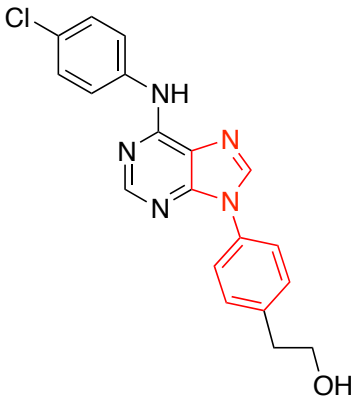
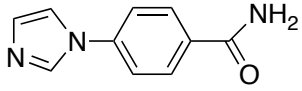
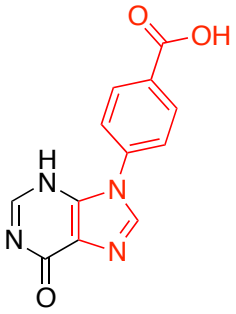
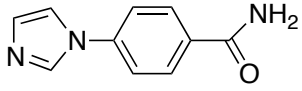
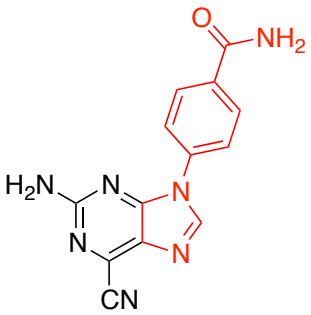
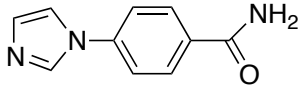
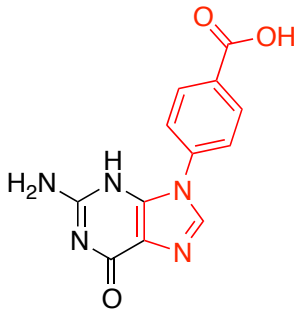
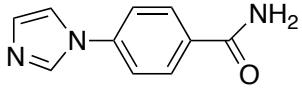
68302890		**F3** 	M.W. 244 H donor: 1 H acceptor: 2 cLogP: 2.14
8895561		**F5** 	M.W. 222 H donor: 1 H acceptor: 2 cLogP: 1.91
7903531		**F5** 	M.W. 203 H donor: 0 H acceptor: 3 cLogP: 1.77
8902266		**F5** 	M.W. 323 H donor: 2 H acceptor: 2 cLogP: 1.77
8906746		**F5** 	M.W. 251 H donor: 1 H acceptor: 1 cLogP: 2.85
6667941		**F5** 	M.W. 258 H donor: 1 H acceptor: 4 cLogP: 1.79

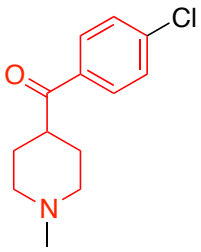
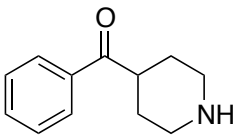
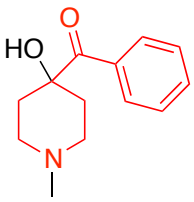
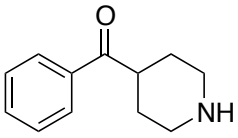
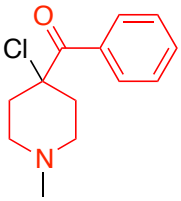
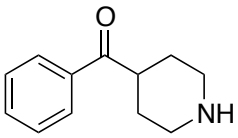
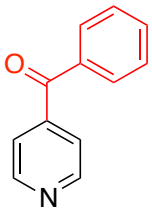
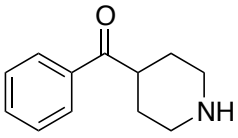
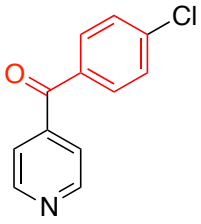
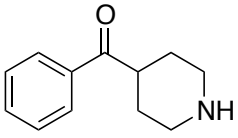
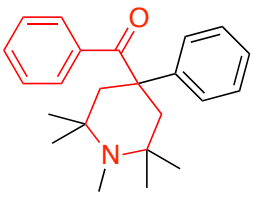
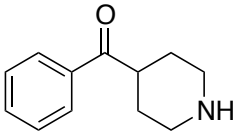
8816655		**F5** 	M.W. 301 H donor: 2 H acceptor: 4 cLogP: 0.47
8902475		**F5** 	M.W. 215 H donor: 1 H acceptor: 1 cLogP: 2.18
8888489		**F5** 	M.W. 237 H donor: 1 H acceptor: 1 cLogP: 2.68
8925662		**F5** 	M.W. 223 H donor: 1 H acceptor: 1 cLogP: 2.41
44815342		**F10** 	M.W. 204 H donor: 1 H acceptor: 1 cLogP: 3.38

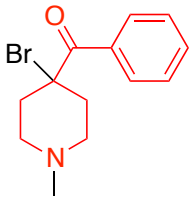
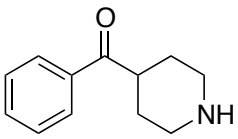
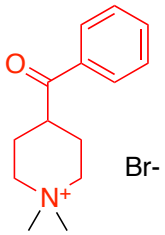
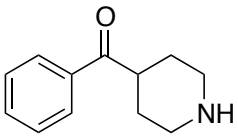
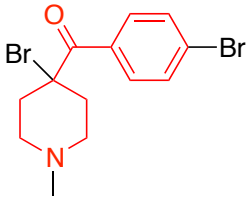
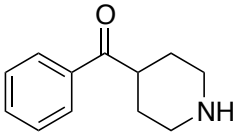
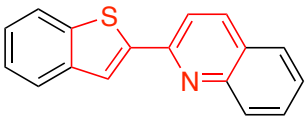
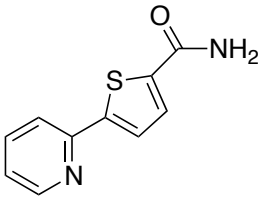
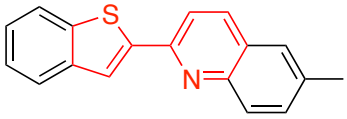
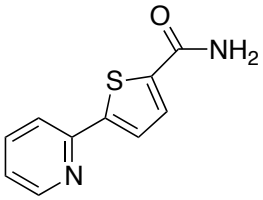
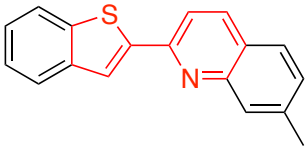
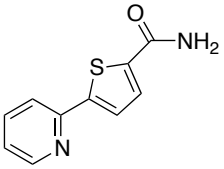
5767021		**F10** 	M.W. 204 H donor: 1 H acceptor: 1 cLogP: 3.38
5149753		**F10** 	M.W. 165 H donor: 1 H acceptor: 2 cLogP: 1.04
5107839		**F10** 	M.W. 199 H donor: 0 H acceptor: 1 cLogP: 3.75
5107788		**F10** 	M.W. 215 H donor: 0 H acceptor: 2 cLogP: 3.54
5108799		**F10** 	M.W. 150 H donor: 1 H acceptor: 1 cLogP: 1.81
5108103		**F10** 	M.W. 165 H donor: 2 H acceptor: 1 cLogP: 2.24

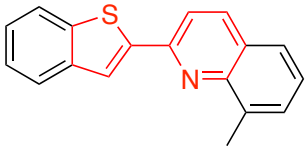
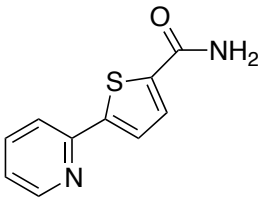
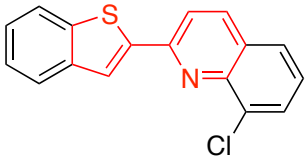
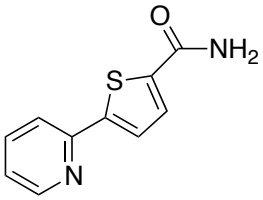
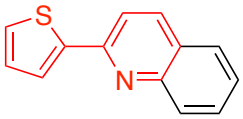
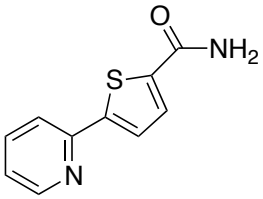
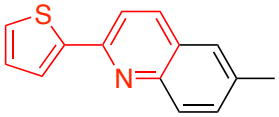
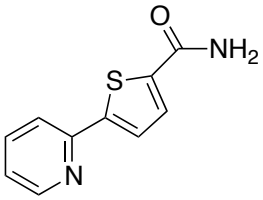
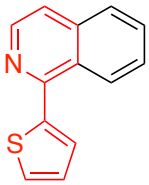
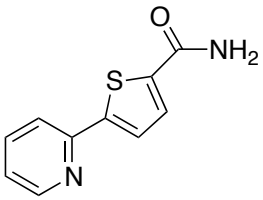
5107833		**F10** 	M.W. 178 H donor: 1 H acceptor: 1 cLogP: 2.76
6060366		**F10** 	M.W. 259 H donor: 1 H acceptor: 1 cLogP: 1.12
NSC82638		**F3** 	M.W. 226.23 H donor: 1 H acceptor: 2 cLogP: 1.5
NSC210362		**F3** 	M.W. 283.33 H donor: 2 H acceptor: 2 cLogP: 2.8
NSC29607		**F3** 	M.W. 277.28 H donor: 2 H acceptor: 2 cLogP: 1.8

NSC82637		**F3** 	M.W. 244.68 H donor: 0 H acceptor: 3 cLogP: 3
NSC82638		**F3** 	M.W. 226.23 H donor: 1 H acceptor: 2 cLogP: 1.5
NSC82639		**F3** 	M.W. 271.23 H donor: 1 H acceptor: 4 cLogP: 1.3
NSC87694		**F3** 	M.W. 270.29 H donor: 1 H acceptor: 5 cLogP: 1.6
NSC100127		F3	M.W. 293.15 H donor: 0 H acceptor: 3 cLogP: 3.5

NSC87695		**F3** 	M.W. 365.82 H donor: 2 H acceptor: 5 cLogP: 3.8
NSC102833		**F3** 	M.W. 256.22 H donor: 2 H acceptor: 4 cLogP: 0.6
NSC210334		**F3** 	M.W. 279.25 H donor: 2 H acceptor: 6 cLogP: 0.3
NSC210437		**F3** 	M.W. 271.23 H donor: 3 H acceptor: 4 cLogP: 0.4

NSC51515		**F7** 	M.W. 237.73 H donor: 0 H acceptor: 2 cLogP: 2.7
NSC280621		**F7** 	M.W. 219.28 H donor: 1 H acceptor: 3 cLogP: 1.3
NSC20220		**F7** 	M.W. 274.19 H donor: 1 H acceptor: 2 cLogP: 2.3
NSC9488		**F7** 	M.W. 183.20 H donor: 0 H acceptor: 2 cLogP: 2
NSC76045		**F7** 	M.W. 217.65 H donor: 0 H acceptor: 2 cLogP: 2.4
NSC363906		**F7** 	M.W. 355.48 H donor: 0 H acceptor: 2 cLogP: 4.8

NSC132804		**F7** 	M.W. 282.17 H donor: 0 H acceptor: 2 cLogP: 2.8
NSC363907		**F7** 	M.W. 298.22 H donor: 0 H acceptor: 2 cLogP:
NSC363768		**F7** 	M.W. 441.98 H donor: 0 H acceptor: 2 cLogP:
NSC40283		**F2** 	M.W. 261.34 H donor: 0 H acceptor: 2 cLogP: 4.9
NSC71686		**F2** 	M.W. 275.36 H donor: 0 H acceptor: 2 cLogP: 5.3
NSC71687		**F2** 	M.W. 275.36 H donor: 0 H acceptor: 2 cLogP: 5.3

NSC71688		**F2** 	M.W. 275.36 H donor: 0 H acceptor: 2 cLogP: 5.3
NSC71689		**F2** 	M.W. 295.78 H donor: 0 H acceptor: 2 cLogP: 5.5
NSC71690		**F2** 	M.W. 211.28 H donor: 0 H acceptor: 2 cLogP: 3.6
NSC71691		**F2** 	M.W. 225.31 H donor: 0 H acceptor: 2 cLogP: 3.9
NSC71692		**F2** 	M.W. 211.28 H donor: 0 H acceptor: 2 cLogP: 3.5

4.2.4. DSF

The ability of fragment analogs (Table 4.1) to shift the BioA T_m was determined using the same experimental procedures for DSF as described for fragment screening (Section 3.2.1). The protein and ligand concentration used were the same.

4.2.5. Crystallography.

The crystallization conditions used to produce BioA crystals were as described in Chapter III (Section 3.2.5.). The diffraction data for the compounds **F10-1**, **F10-2**, and **F5-1** were collected at 100 K using synchrotron radiation with a Dectris Pilatus 6M Pixel Detector on beamline 17-ID (IMCA-CAT) at APS, Chicago, United States. The data were processed, integrated, and scaled with XDS and SCALA[78] using the autoPROC scripts available at IMCA-CAT. Data for a compound **F10-3** co-crystal were collected at 100 K using Cu *Kα* radiation on a Rigaku HighFlux HomeLab rotating-anode system with a Saturn 944+ CDD detector in the Kahlert Structural Biology Laboratory at the University of Minnesota. This data was processed, integrated and scaled with d*TREK[111]. The data collection and processing statistics are given in Table 4.3. The structures were solved by molecular replacement as described in Chapter III (Section 3.2.5).

4.2.6. ITC

The ITC experimental procedure for compound **F5-1** was as described in Chapter III (Section 3.2.4). ITC experiments were performed by Dr. Feng Liu in the lab of Prof. Courtney Aldrich.

4.2.7. UV-Vis spectroscopy

A NanoDrop 1000 UV–Vis spectrophotometer (Thermo Scientific) was used for all UV–Vis spectroscopy. *Mtb* BioA protein (2.0 µL, 0.16 mM) in HEPES (25 mM, pH 7.5), NaCl

(50 mM, 1 mM EDTA), TCEP (0.1 mM) was mixed with compound **F10-1** (2.0 µL, 0.4 mM) in HEPES (25 mM, pH 7.5), NaCl (50 mM), EDTA (1 mM), TCEP (0.1 mM) and its UV-Vis spectrum was immediately measured upon mixing. UV-Vis spectra were taken at 0 s, 15 s, 30 s, 60 s and 120 s after mixing.

4.2.8. Steady-state Kinetics

Mode of inhibition studies were carried out under initial velocity conditions in a total volume of 50 µL at 25 °C in 384 well black plates (Corning 3575). Reactions were set up in triplicate and consisted of BioA (114 nM) in reaction buffer (100 mM Bicine pH 8.6, 50 mM NaHCO3, 1 mM MgCl2, 0.0025% Igepal CA-630, 5 mM ATP, 0.1 mM PLP, 320 nM E. coli BioD, 20 nM Fl-DTB, 184 nM streptavidin, and 1 mM TCEP) with either variable amounts of KAPA (0.94–7.5 µM) with SAM (2.34 mM) or with a variable amount of SAM (0.3–2.5 mM) with a fixed amount of KAPA (1.9 µM). Each substrate concentration was run with 0, 31.25, 62.5, and 125 µM inhibitor. Reactions were monitored on a microplate reader using an excitation of 485 nm, and emission at 535 nm. A standard curve of dethiobiotin (2.7 nM–2 µM) in reaction conditions lacking only BioA was used to convert fluorescence into enzyme velocities as previously described[73]. The data was fit using the enzyme kinetics module of SigmaPlot to competitive, uncompetitive, and non-competitive models and the model with the highest r2 value was selected. The PLP utilizing enzymes alanine transaminase (ALT, EC 2.6.1.2) and aspartate transaminase (AST, EC 2.6.1.1) were used to test off target inhibition of **F10-1**. Reactions were carried out in 100 µL in 96 well UV clear half-area plates (Corning 3679). For ALT, reactions consisted of enzyme (10 mU) in reaction buffer (60 mM Bicine pH 8.0, 0.1 mM NADH, 1 mM TCEP) containing 100 mU lactate dehydrogenase (LDH) and either fixed α-ketoglutarate (75 µM) with 0.625–10 mM alanine or fixed alanine (10 mM) with 6.25–50 µM α-ketoglutarate. Each concentration of substrate was run with DMSO

or 29.6–66.7 μM **F10-1**. The reactions (in duplicate) were monitored by the decrease in A340 that corresponds to the consumption of NADH by LDH upon the formation of pyruvate from ALT. Initial velocities were calculated using the molar absorptivity of NADH (6220 M-1 cm-1 at 340 nm) and fit as described above using SigmaPlot. To test for inhibition of AST 2 mU of enzyme in reaction buffer containing 100 mU malate dehydrogenase (MDH), 236 μM α-ketoglutarate and 0.156–2.5 mM aspartate was tested against **F10-1** (100 μM). The kinetic assay was done by Mr. Daniel Wilson in the lab of Prof. Courtney Aldrich.

4.3 Results

4.3.1. Substructure search, similarity search and structure based virtual screening

50 commercially available compounds were selected to build the initial library from fragment optimization. These 50 compounds all passed through the Lipinski Rule of 5 filter. We superimposed these optimized compounds in the fragment position in the 3D complex structures. By excluding molecules that had steric conflicts with the binding pocket, the number of compounds for purchase was reduced to 16 (Table 4.2).

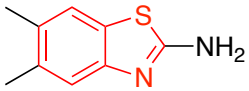
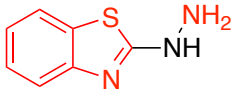
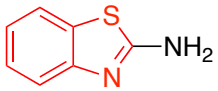
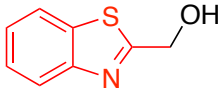
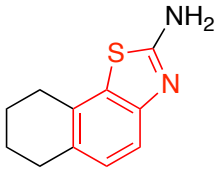
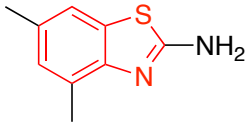
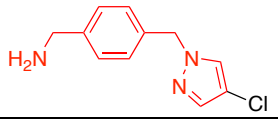
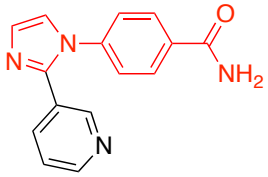
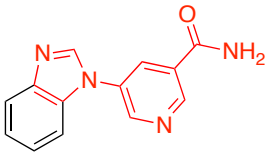
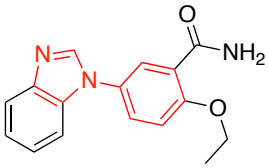
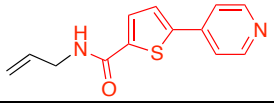
4.3.2. DSF and X-ray crystallography of compounds from commerce

DSF was used to measure the T_m shifts of BioA upon binding with the 16 purchased compounds. Detailed T_m shifts are listed in Table 4.2. Nine out of sixteen compounds induced a T_m shift equal or greater than ± 2 °C. We used both soaking and co-crystallization to try to obtain complex X-ray structures. Co-crystallization techniques resulted in better crystal quality and higher ligand occupancy than the soaking method. As an additive to the reservoir solution, the adduct-forming compounds increased the rate of complex crystal growth. The co-crystals have potential to diffract to beyond 1.3 Å

as we observed in data collection. Complex structures were obtained for 2 of the 16 compounds purchased. These are identified as **F10-1** and **F5-1**, (Figure 4.2).

Table 4.2 Biophysical Characterization Of Compounds From Commerce.

The fragment core structures maintained in the designed compounds are colored red.

Cmpd ID	Structure	DSF T_m(°C)	ΔT_m (°C)	Fragment Template	X-ray data
5107833		83	-2	**F10**	
F10-1 5108103		67	-18	**F10**	Adduct Complex
5108799		83	-2	**F10**	
5149753		83	-2	**F10**	
5767021		83	-2	**F10**	
6060366		73	-12	**F10**	
F5-1 8895561		74	-11	**F5**	Complex
10896690		87	+2	**F3**	
13846515		78	-7	**F3**	
94450964		85	0	**F3**	
68302890		85	0	**F2**	

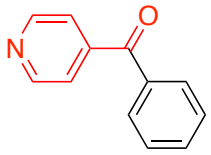
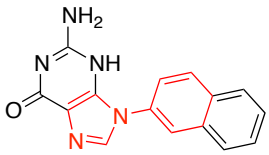
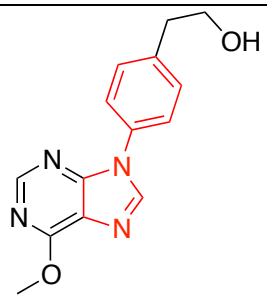
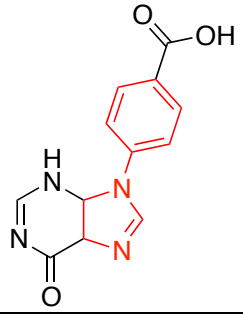
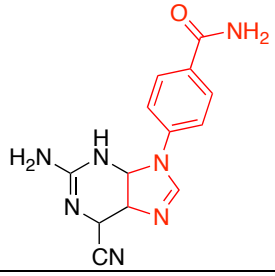
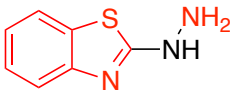
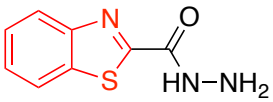
NSC9488		85	0	**F7**	
NSC29607		85	0	**F3**	
NSC87694		85	0	**F3**	
NSC102833		85	0	**F3**	
NSC210334		85	0	**F3**	

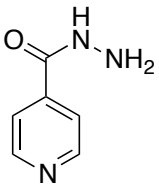
Figure 4.2 Acquired fragment analogs yielding new crystallographic complexes.



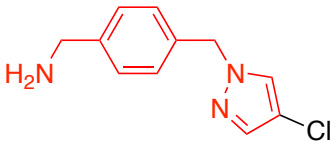
F10-1
H2L12 T_m 67°C



F10-2
H2L26 T_m 70°C



F10-3 (Isoniazid)
H2L43 T_m 71°C



F5-1
H2L17 T_m 74°C

The fragment core structures maintained in the designed compounds are colored red.

4.3.3. **F5** optimization

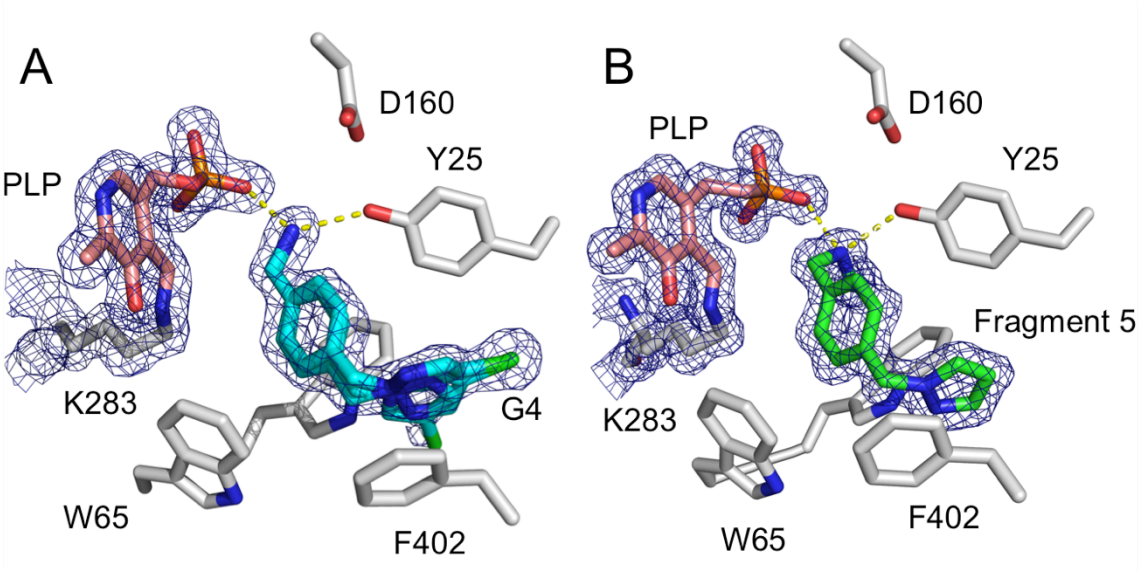
Compound **F5-1** was selected because from the **F5** complex structure it appeared that replacing the terminal secondary amine with a primary amine would result in more favorable configuration for hydrogen bonding interactions with the PLP phosphate oxygen and the Tyr25 hydroxyl group (Figure 4.2). **F5** induced a stabilizing shift of four degrees (T_m = 89 °C), while compound **F5-1** induced a destabilizing shift of eleven degrees (T_m = 74 °C) (Figure 4.2).

The crystal structure of the complex with **F5-1** has been determined at 1.62 Å resolution (Table 4.3) and is illustrated in Figure 4.3A. The compound binds in the same site as **F5** and in a very similar conformation (Figure 4.3). The conformations of the key residues (Y25, W64, W65, F402) in BioA active site are also the same as the **F5** complex. **F5-1** differs in two respects from **F5**. First, **F5-1** has a terminal primary amine while **F5** has a secondary amine. The primary amine has a more favorable electronic configuration to form hydrogen bonding interactions with both PLP phosphate oxygen and Tyr25 hydroxyl group. On the other hand however, the secondary amine has the terminal methyl group as an electron pushing group that should make the secondary amine a stronger electron donor in the H bonding. With this analog, we were interested to see if the terminal methyl group is necessary for the fragment binding, and found that it is not crucial. Compound **F5-1** also includes a chloride substituent on the pyrazole ring not present on the parent fragment. The electron density for the ligand in the **F5-1** complex shows that the pyrazole chloride can exist in at least three different conformations that arise from alternate rotations of the pyrazole ring about the linking bond. The nitrogen atoms in the pyrazole ring make no specific hydrogen bonding interactions with the surrounding residues that might make one conformation more favorable than another.

The key BioA amino acid residues in the binding site retain the same conformations as observed in the complex with **F5**.

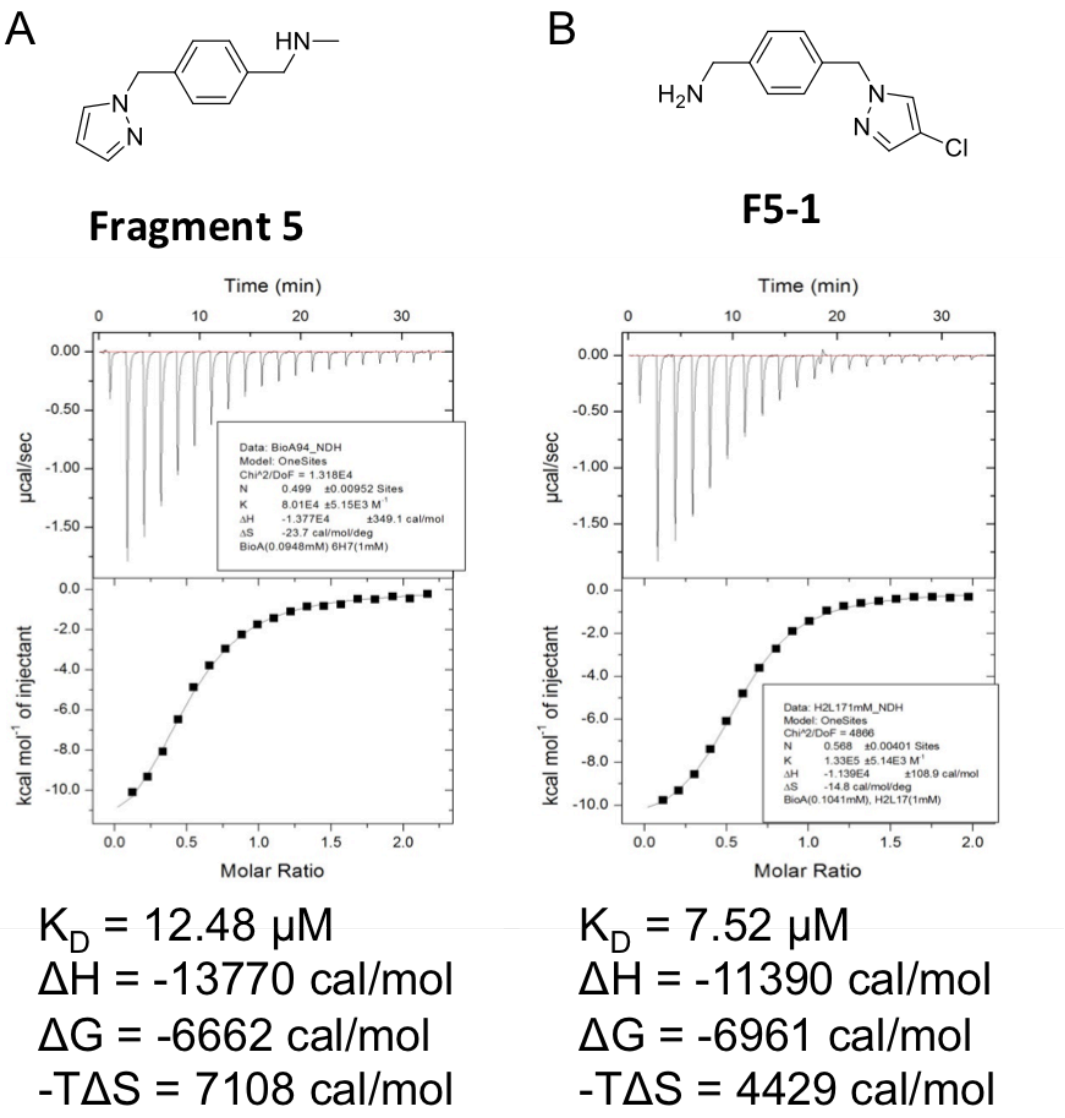
ITC was used to characterize the binding affinity of compound **F5-1**. The K_D value is 7.5μM (Figure 4.4A), which is modestly (1.7 fold) better than the K_D of **F5** (12.5 μM) (Figure 4.4B). This result agrees with our structure observation from X-ray crystallography. The modification of the secondary amine to a primary amine maintained similar strength of the hydrogen bonding interactions; **F5** has a ΔH of binding of −13.7 kcal/mol while **F5-1** has a ΔH of −11.4kcal/mol (Figure 4.4).

Figure 4.3. A comparison of BioA complexes with Compounds F5-1 and F5.



(A) Complex with compound **F5-1**, and (B) with **F5**. 1σ $2F_o$-F_c electron density for the ligand, Lys383, and the PLP is shown for each complex (mesh).

Figure 4.4 ITC results for F5 and its optimized analog F5-1.



ITC results for (A) **F5**, and (B) **F5-1**. The upper figures show the time dependence of the electric power (µcal/sec) to maintain constant temperature of the sample after each injection. The lower figures show the heat per mole of injectant vs. the molar ratio of the ligands and protein in the system.

4.3.4. Discovery of a reversible covalent inhibitor from **F10** optimization

Compound **F10-1** was selected as a promising analog because from the **F10** complex structure it appeared that a nitrogen in the place of the **F10** α-carbon could provide an extra hydrogen bonding interaction with Thr318′ hydroxyl group. This compound induced an even larger destabilizing T_m shift of nineteen degrees (T_m = 67 °C) (Figure 4.2). It did not escape our attention that this is nearly the T_m assigned to *apo* BioA.[74]

As crystal soaking experiments were begun with Compound **F10-1**, it became clear that this analog behaves differently in the presence of BioA. The PLP-bound *holo* enzyme crystals typically grown under the same conditions are yellowish in color.[25] *Holo* BioA crystals change color from yellow to red within 2 minutes when placed in the **F10-1** soaking solution (Figure 4.5). In co-crystallization experiments, a similar phenomenon was observed. Within seconds of the addition of Compound **F10-1** to crystallization drops, the solution develops a deep red color suggestive of chemical reaction. The reaction of **F10-1** with BioA can be followed spectroscopically, with the time-dependent increase in absorbance at 330 and 500 nm and a decrease at 420 nm that suggests the reaction is complete in just a few minutes (Figure 4.5). Co-crystals of the complex were prepared under the same conditions as *holo* crystals, but the red color clearly concentrates in the crystals. Diffraction data for Compound **F10-1** co-crystals were collected to 1.90 Å resolution (Table 4.3).

The crystal structure of the complex with **F10-1** is shown in Figure 4.6. The hydrazine analog forms a covalent adduct with the PLP aldehyde to form an extended *cis*-azo quinonoid species. The bond between PLP and Lys283 is clearly broken, and the lysine amino group is relocated to H-bond with Thr318′. This was clearly not the case with bound fragment **F10**, where the bond between Lys-283 and the PLP is intact (Figure 3.2. F). A planar Schiff base is formed with the hydrazine, and the entire benzothiazole is well-defined in unambiguous electron density (Figure 4.6A). While the benzothiazole can still be said to occupy the SAM binding subsite, it is rotated ~90° from the position occupied by its template, **F10**, shifted significantly away from Tyr 25, and out of direct contact with Phe402 (Figure 4.6A). Two localized enzyme conformational adjustments involving only the reorientation of side chains occur in conjunction with adduct formation. The hydroxyl group of Tyr 407 shifts 2.5 Å from the *holo* enzyme position to make a new H-bond to the carbonyl oxygen of Arg400. If unchanged, a short contact to benzothiazole C7 would have existed. In the largest conformational change, the side chain of Trp65 rotates from the holo position (chi1=-60°; chi2=-20°) to a new position (chi1=-60°; chi2=90°) so that the hydrogen on the indole Nε1 is positioned for optimal interaction with the benzothiazole π system. Tyr25 occupies two different positions in this complex with roughly equal occupancy. In one conformation, the OH is H-bonded to Asp160 and the OH of Tyr157 as in the *holo* structure. In the other, the side chain is rotated so that the OH makes an H-bond with the alternate carboxylate oxygen of Asp-160 (Figure 4.6A).

To determine whether or not the inhibition of BioA by Compound **F10-1** is reversible, the BioA–Compound **F10-1** adduct was dialyzed, which resulted in loss of the deep red color and full restoration of the enzyme activity. To further characterize the modality of inhibition, steady-state kinetic studies were performed under initial velocity conditions.

Double reciprocal plots of initial reaction velocity under conditions with varying inhibitor and reactant concentration (Figure 4.7A) show that Compound **F10-1** is a competitive reversible inhibitor of BioA with respect to SAM, (K_i = 10.4 ± 0.6 μM), and an uncompetitive inhibitor with respect to KAPA (K_{iu} = 85.4 ± 3.4 μM). This is the expected behavior for an inhibitor of one step of a ping pong bi-bi mechanism[97]. In forming a reversible adduct with the PLP cofactor, Compound **F10-1** inhibits the same BioA enzyme form (PLP-BioA) with which SAM reacts. The inhibition pattern against KAPA is uncompetitive, because KAPA binding is required to regenerate the PLP-bound form of BioA.

In an effort to explore whether Compound **F10-1** possesses any selectivity, it was evaluated as an inhibitor of alanine transaminase (ALT), and of aspartate transaminase (AST), two other important mammalian PLP-dependent transaminases[98]. Mode of inhibition studies reveal that Compound **F10-1** is an uncompetitive reversible inhibitor of ALT with respect to α-ketoglutarate with a K_{iu} value of 74.0 ± 8.2 μM and a competitive inhibitor with respect to alanine (K_i = 18.9 ± 1.6 μM) (Figure 4.7B). As with BioA, the inhibitor competes for binding with the compound that is the amino group donor responsible for converting the BioA protein from PLP bound holo form to the pyridoxamine phosphate (PMP) state. We did not observe any detectable inhibition of AST, however (data not shown). While Compound **F10-1** may not be a selective inhibitor of BioA, it is also not entirely non-specific.

Armed with the knowledge that Compound **F10-1** is a reversible competitive inhibitor of BioA, we chose to investigate attributes of the analogous hydrazide (Compound **F10-2**; Figure 4.2). Compound **F10-2** also gave rise to a large destabilizing T_m shift (-15 °C;

Figure 4.2). We were also able to characterize the complex by co-crystallization (Table 4.3). No visible color change is observed upon addition of **F10-2** to BioA.

The crystal structure of Compound **F10-2** also confirms the formation of a covalent adduct, but again, the bound adduct conformation is unique (Figure 4.6B). From high resolution diffraction data (1.7 Å resolution) we can assert that conjugation throughout the Schiff base appears complete; each of the atoms is sp^2-hybridized and planar in its bonding, but there is a subtle curvature imposed throughout the adduct that likely implies some degree of electronic strain. While the side chain of Trp65 remains primarily in the rotated position required to accommodate the adduct with **F10-1**, the benzothiazole in the complex with **F10-2** is positioned well above and out of contact with it, but stacks instead against the indole of Trp64. **F10-2** induces a side chain configuration similar to that induced by **F7** (Chapter III Table 3.4), except that Y25 adopted the conformation in **F3** complex. The hydrazide carbonyl is oriented toward the Trp65 indole nitrogen, but it is 3.9 Å away from it – too far to be considered a stabilizing hydrogen bond. A water-mediated interaction is possible, but no well-ordered water molecule is observed. The benzothiazole also does not point directly toward Phe402, but instead reaches past it and into the larger pocket bounded by Arg400, the side chain that interacts directly with the carboxylate of KAPA. Tyr25 is observed in the same two conformations found in the complex with **F10-1**.

The structural similarity between Compound **F10-2** and isoniazid (Compound **F10-3** in Figure 4.2), one of several chemical agents that comprise the cocktail of drugs often used to treat *Mtb* infections, led us to evaluate this aryl hydrazide as a BioA inhibitor. In the same battery of biophysical studies, isoniazid appears to interact with BioA. It produces a more modest but still large negative shift in the T_m (-8 °C), and crystallographic analysis shows that it forms a covalent adduct with the PLP that is very

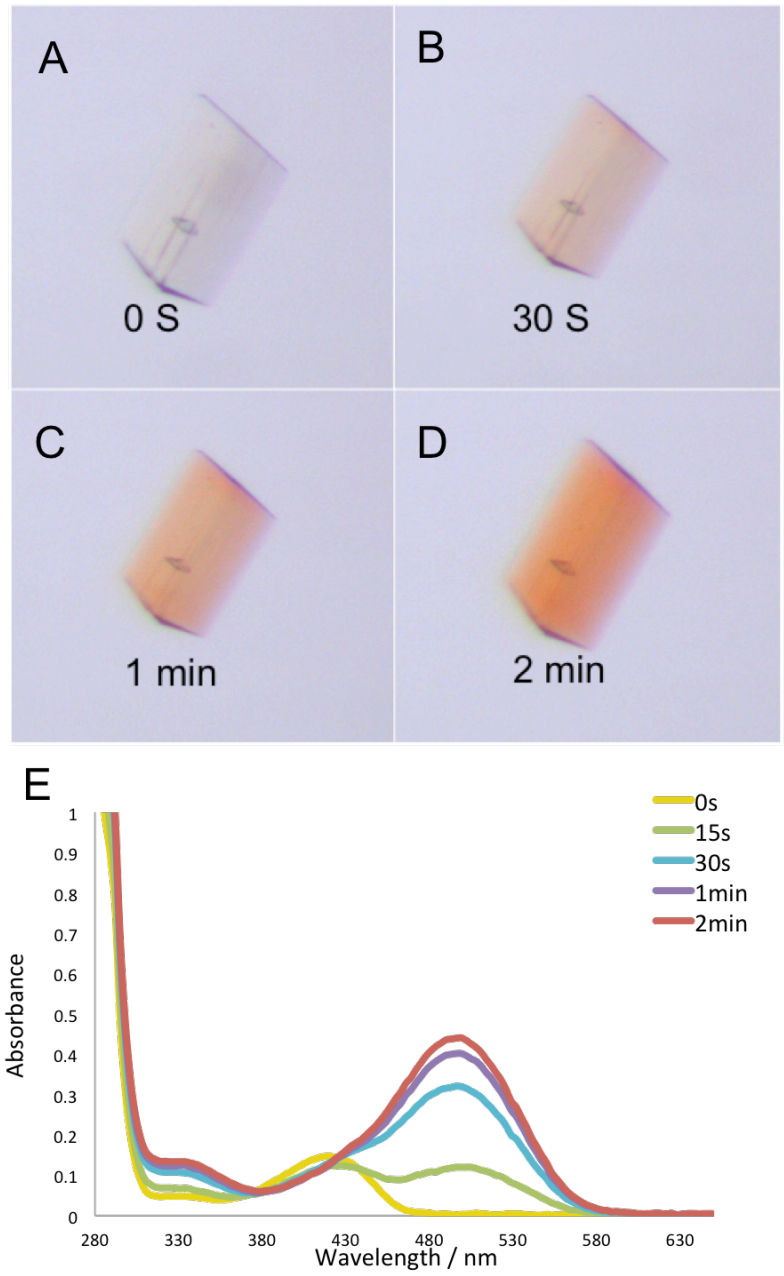
similar to that formed by compound **F10-2** (Figure 4.6C). Nevertheless, in the coupled BioA assay used to assess inhibition kinetics[73], both hydrazides **F10-2** and **F10-3** show no inhibitory activity toward BioA (K_i > 100 mM). We are led to conclude that these adducts seen crystallographically are so easily reversed upon exposure to substrates that no significant amount of competitive inhibition can be observed by these low affinity compounds.

Table 4.3 Collection and refinement statistics of the optimized compounds F5-1, F10-1, F10-2 and F10-3

Ligand ID	F5-1	F10-1	F10-2	F10-3 Isoniazid
PDB id	4WYG	4MQP	4MQQ	4MQR
Data collection site	IMCAT	IMCAT	IMCAT	in house
Detector	DECTRIS PILATUS 6M	DECTRIS PILATUS 6M	DECTRIS PILATUS 6M	Saturn 944+ CCD
Wavelength (Å)	1.000	1.000	1.000	1.541
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions				
a, b, c (Å)	63.12, 66.43, 205.32	63.02, 65.92, 201.96	62.95, 66.35, 203.97	63.13, 66.48, 203.68
α,β,γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
molecules per ASU	2	2	2	2
Resolution (Å)	205.32-1.62 (1.625-1.62)	201.96-1.83 (1.93-1.93)	63.05-1.55 (1.555-1.55)	63.20-2.10 (2.18-2.10)
R_{merge}	0.058(0.190)	0.063(0.284)	0.088(0.366)	0.055(0.182)
I/σ_I	22.3(8.7)	18.0(5.5)	14.0(2.7)	16.6(2.8)
Completeness	96.2%(94.7%)	98.3%(96.6%)	99.0%(86.0%)	88.4%(73.4%)
Multiplicity	6.8(6.4)	6.4/6.3	6.2/4.5	3.22(2.17)
No. observations	725282	467939	757129	145139
No. unique reflections	106327	73075	121718	45088
Refinement				
Resolution (Å)	46.40-1.62	100.93-1.83	101.99-1.70	101.84-2.10
Rwork/Rfree	15.2/18.0	17.8/20.8	17.6/20.4	20.7/25.5
No. atoms	7951	7106	7597	6885
No. water	879	474	747	375
No. ligand molecule	2	2*	2*	2*
No. PLP molecule	2	2	2	2
No. other molecule	11	4	4	4
Ramachandran plot				
Favored	706 (96.4%)	758 (95.0%)	707 (96.7%)	784 (94.0%)
Allowed	22(2.9%)	27(3.4%)	20(2.7%)	35 (4.2%)
Disallowed	5(0.6%)	13(1.6%)	4(0.6%)	15 (1.8%)
R.m.s deviations				
Bond lengths (Å)	0.008	0.009	0.009	0.01
Bond angles (°)	1.15	1.34	1.27	1.31

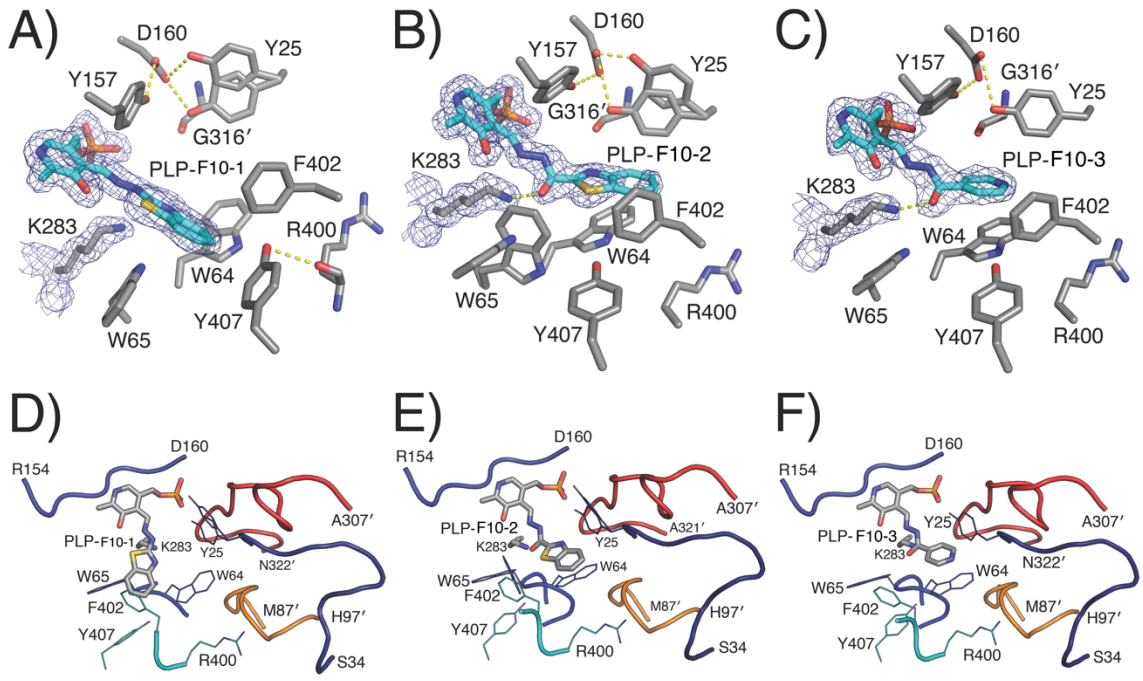
* PLP molecules are in adduct form with ligands in the structure.

Figure 4.5. Colorimetric Characterization of F10-1 binding.



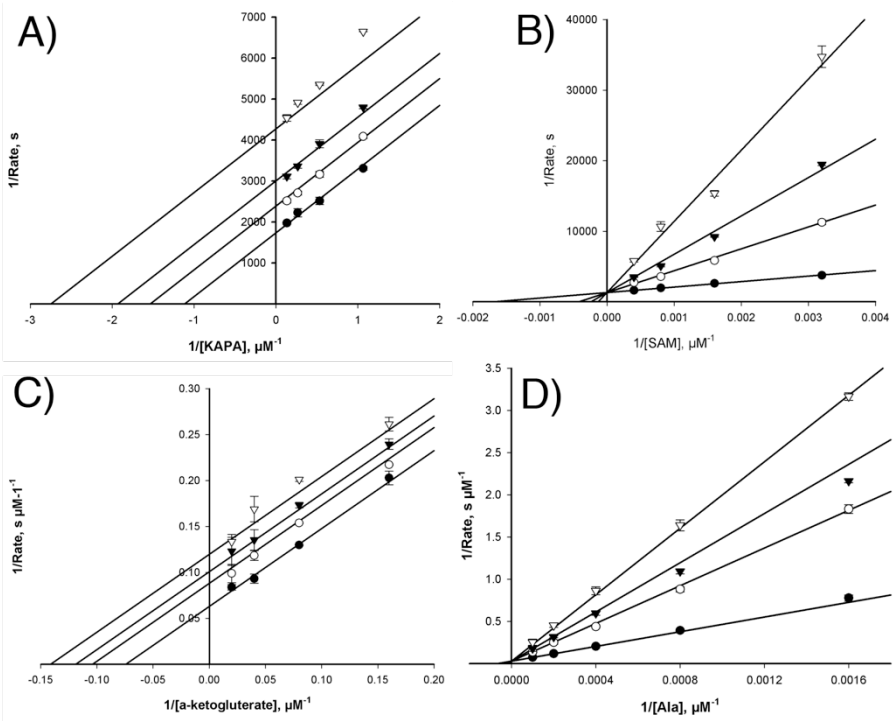
A-**D**: Time-lapsed photos of a PLP-bound BioA crystal soaked in Compound **F10-1** (15% PEG 8000, 100 mM HEPES pH 7.5, 100 mM $MgCl_2$, and 5 mM Compound **F10-1**) over 2 minutes. **E**: UV-Vis spectroscopy of 0.16 mM PLP bound BioA (Holo BioA) upon mixing with 0.4 mM Compound **F10-1** at different time points.

Figure 4.6. Crystal Structures of analogs of F10.



A-C) A comparison of complexes of covalent adducts created by reaction of PLP with Compounds **F10-1**, **F10-2**, and **F10-3**. Covalent adducts are labeled PLP-**F10-1**, PLP-**F10-2**, and PLP-**F10-3**, respectively. 1σ $2F_o$-F_c electron density for Lys383, the PLP and each ligand is shown (mesh). Electron density clearly shows that the bond to Lys283 is replaced with a covalent bond to each compound. Tyr25 is shown in two conformations when so observed. Local binding environment of each adduct (**F10-1-F10-3**) is shown in the same orientation in frames D-F) to emphasize structural similarities and differences.

Figure 4.7 Inhibition studies with Compound F10-1



Initial rate data of variable amounts of inhibitor and either KAPA (**A**) or SAM (**B**). KAPA was varied from 0.94 to 7.5 μM and SAM was varied from 0.3–2.5 mM with the fixed substrates held at 2.34 mM for SAM and 1.9 μM for KAPA. **F10.1** was used at concentrations of 0 μM (●), 31.3 μM (○), 62.5 μM (▼), 125 μM (▽). The inhibitor is uncompetitive with respect to KAPA and competitive with respect to SAM with K_i values of 85.4 ± 3.4 μM and 10.4 ± 0.6 μM respectively. Inhibition of ALT by **F10.1**. **C**)-**D**)**:** Initial rate data of variable amounts of inhibitor with respect to α-ketoglutarate (**C**) or alanine (**D**). The substrate α-ketoglutarate was varied from 6.25 to 50 μM and alanine was varied from 0.625 to 10 mM with the fixed substrates held at 10 mM for alanine or 75 μM for α-ketoglutarate. Compound **2** was used at concentrations of 0 μM (●), 29.6 μM (○), 44.4 μM (▼), and 66.6 μM (▽). The inhibitor is uncompetitive with respect to α-ketoglutarate and competitive with respect to alanine with K_i values of 74.0 ± 8.2 μM and 18.9 ± 1.6 μM, respectively.

4.4 Discussion

4.4.1. Stabilization of an *azo*-quinonoid intermediate

Compounds **F10-1**, **F10-2** and **F10-3** all react with the PLP co-factor reversibly to form stable covalent adducts. The UV-Vis spectroscopy and crystallography results for Compound **F10-1** confirm that the hydrazine analog forms an extensively conjugated *cis*-azo-quinonoid species. While similar quinonoids form transiently as intermediates in the catalytic mechanism of all transaminases, they are rarely captured.[99] We propose a mechanism for the stabilization of this adduct from **F10-1** in Figure 4.8A. By analogy to the reaction with substrates, the hydrazine first displaces the lysine to form a PLP-coupled imine. Tautomerization of this imine leads to the formation of an azo-quinonoid; the delocalized bonding is responsible for the beautiful sanguine color. Further reversible tautomerization of the pyridoxal ring is possible through conjugation all the way to the benzothiophene (Figure 4.8B). This is a unique chemical feature of Compound **F10-1** not shared with the hydrazides (**F10-2** or **F10-3**) that can only form less extensively delocalized imines (Figure 4.8C). While these hydrazide adducts are evidently more stable than the PLP-lysine Schiff base that they supplant, the fact that these compounds show no detectable inhibition in competition with substrates suggests that their stability is low. The hydrazine adduct, however, with a K_i of 10 μM is unusually stable, which is approximately 80-fold lower than the K_m of SAM.[72]

The reactivity of hydrazines and hydrazides with PLP dependent transaminases has been known for decades and the kinetics of inhibition were studied in detail long ago.[100] More recently, there have been a number of attempts to increase the potency of

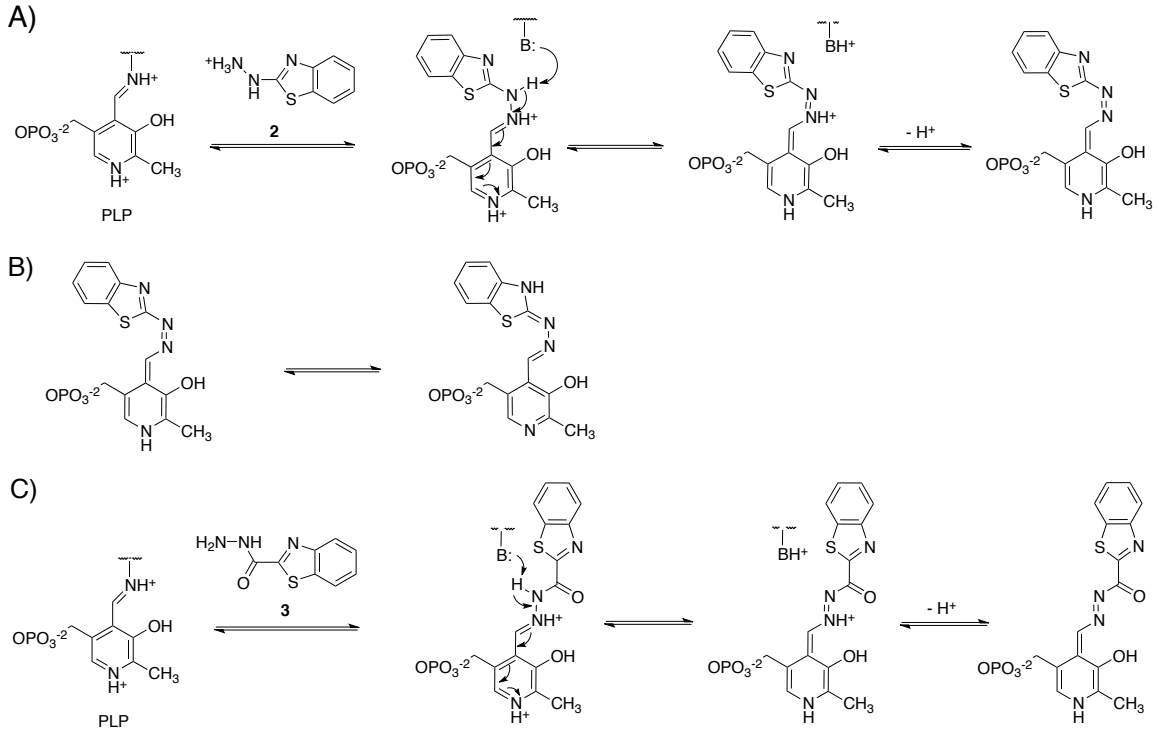
inhibitors of PLP-dependent enzymes that contain a reactive hydrazide. Ejim *et al*. confirmed the formation of a stable adducts to PLP bound in cystathionine beta-lyase with a series of hydrazinocarbonylmethylbenzamides, and were able to improve inhibitor binding affinity by 50-fold with the preparation of small analog library.[101] Another hydrazide has recently been identified as an inhibitor of *E.coli* BioA by whole-cell phenotypic screening.[102] Others have investigated a series of aryl hydrazides as inhibitors of LL-diaminopimelate aminotransferase. Structure-activity relationships revealed in the study of these systems have identified the reactive hydrazide as a necessary, but not sufficient molecular feature needed for inhibition. Potent inhibition can only occur when other binding site complementarity also exists. These examples do not benefit from the extensive electronic delocalization that occurs upon reaction with the hydrazylbenzothiophene of Compound **F10-1**. For this reason, we suggest that this compound may serve as a particularly effective starting point for further optimization. Whether selective inhibitors of BioA can be generated from the 2-(hydrazinyl)benzo-thiophene lead remains to be determined, but the complexes with compounds characterized as part of this study have exposed a variety of different conformational states that are accessible to the BioA enzyme that should prove valuable in the structure-based design of more potent inhibitors based on this scaffold. Our analysis of the hydrazides and hydrazines derived from optimization of **F10** has been recently published[62].

4.4.2. The Value Of Destabilizing Fragment Hits

Many of the studies evaluating the use of thermal shift methods for identification of protein ligands focus only on molecules that increase the transition temperature.[82, 103-104] All of the new compounds that were discussed in this chapter shift T_m significantly *downward*. This is true of the initial non-covalent template for **F10-1**, **F10-2** and **F10-3**

(**F10**, ΔT_m = -5 °C), as well as the reversible adducts with ΔT_m shifts in the range of -6 to -16 °C. Although the template for Compound **F5-1** (**F5**) induces an upward T_m shift (ΔT_m = 4 °C), **F5-1** has a ΔT_m shift of -11 ^{o}C. In prior DSF studies, we have tentatively ascribed a much lower T_m (67 °C) to the apo (PLP-free) form of BioA,[74] and were initially drawn to these fragment hits because we expected to find that the molecules bound in place of the PLP. Structural characterization of these compounds clearly shows this is not the case. Why these compounds are so sharply destabilizing is not known; it may be that they play a greater role in the stabilization of the unfolded state. Empirically, it is useful to note for the benefit of those hoping to conduct similar studies with other proteins, that destabilizing compounds may also be worth structural investigation. Compounds that destabilize a protein are expected to lead to more rapid protein degradation. In *Mtb,* damaged proteins are removed by the Pup-proteasome system. Thus small molecules that lead to protein destabilization could potentially result in protein depletion, which may be advantageous under non-replicating conditions where protein synthesis is not occurring to replenish degraded proteins.

Figure 4.8. Mechanisms of covalent adducts of hydrazines and hydrazides



A) Proposed mechanism of **F10-1**-PLP adduct formation; **B**) Tautomerization of **F10-1**-PLP adduct; **C**) Proposed mechanism of **F10-2**-PLP adduct formation.

4.5. Conclusion

In this chapter, we have described our work toward the structure-aided optimization of six BioA fragment hits. Commercial based SAR and binding structure based virtual screening allowed us to experimentally work on a manageable number of purchased compounds. We have identified 4 ligands that can be co-crystallized with BioA out of a total of 19 compounds we purchased; 2 of them were confirmed with improved binding affinity or inhibition (**F5-1** and **F10-1**).

Chapter V: Fragment structure-based optimization of BioA HTS hits

Acknowledgement: Sections 5.2.4-5.2.6, 5.3.2 first paragraph and figure 5.2 include excerpts from the work [S.W. Park, et al., "Target-Based Identification of Whole-Cell Active Inhibitors of Biotin Biosynthesis in Mycobacterium tuberculosis". Chem. Biol.In Press.]. Prof. Barry Finzel contributed the writing of section 5.3.2 first paragraph and making figure 5.2 (A-C). The ITC experiments in this section were performed by Dr. Feng Liu in the lab of Prof. Courtney Aldrich.

5.1 Introduction

Protein-fragment complex structures provide invaluable information for rational inhibitor design. As discussed in chapter I, three different strategies are often used in fragment based lead design. Fragment extension is used to grow a single fragment into adjacent sub-sites; and fragment linking can be applied when two fragments bind in two sub-sites and they can be tethered using a linker. In the real world, however, there may be multiple fragments or lead compounds that bind in the same site of the target but make different interactions[105]; when these bound compounds partially overlap with each other, fragment merging strategies can be used to optimize the lead compounds.

Fragment merging strategies are not limited to fragments; the same approach can be used to optimize more potent lead compounds whenever diverse molecules share the same binding site. This strategic recombination of ligand molecular features has often been called “scaffold hopping”. It requires a structural understanding of how lead compounds bind to the targets from experimental or virtual models. It aims at interchanging core ligand structural features (the scaffold) of lead compounds in a way

that allows them to retain a predictable binding conformation and biological potency[106]. The purpose of the "hopping" to a different scaffold is to find novel active chemotypes that are also "drug-like". Scaffold hopping is often used to optimize the pharmacokinetic properties of lead molecules (absorption, distribution, metabolism, and excretion, profiles), or to achieve improved selectivity toward the intended target.

Scaffold hopping is routinely used to optimize lead compounds from HTS when structural knowledge of the binding of multiple compounds is known. Traditional scaffold hopping method includes heterocycle replacements, ring opening or closure, peptidomimetics, topology-based hopping and small-step hopping[107]. Virtual tools and computational models may be used to assist in design to reduce the high expense required to synthesize unpromising compounds[108]. Still, computational designs can be better supported if more experimental structural information is available. The beauty of a fragment merging strategy for optimization of HTS leads is that, by combining varied fragment binding features with known lead binding features, the effective number of scaffolds available for recombination attempts can be greatly expanded, thereby improving the likelihood that improved molecules can be designed.

Recently Baeschlin *et al.* [109] described such an excellent example. In their work, structure information from fragment complex structures obtained using X-ray crystallography were combined with other structures to identify a higher affinity inhibitor of renin. In their research, an NMR-based screen was applied to a small focused library of 113 compounds against the aspartyl protease. Hit compound 1 (Figure 5.1) identified by fragment screening was crystallized with the target and its binding features were studied. In parallel, a HTS screening effort resulted in the discovery of lead compound 2 (Figure 5.1). Crystallographic structures showed that the diphenylmethane moiety of compound 2 bound in a similar position as the tricycle of compound 1. Upon

incorporating the tricyclic framework into compound 2, they derived compound 3 with enhanced potency. Further optimization led to discovery of compound 4 with nanomolar potency (Figure 5.1).

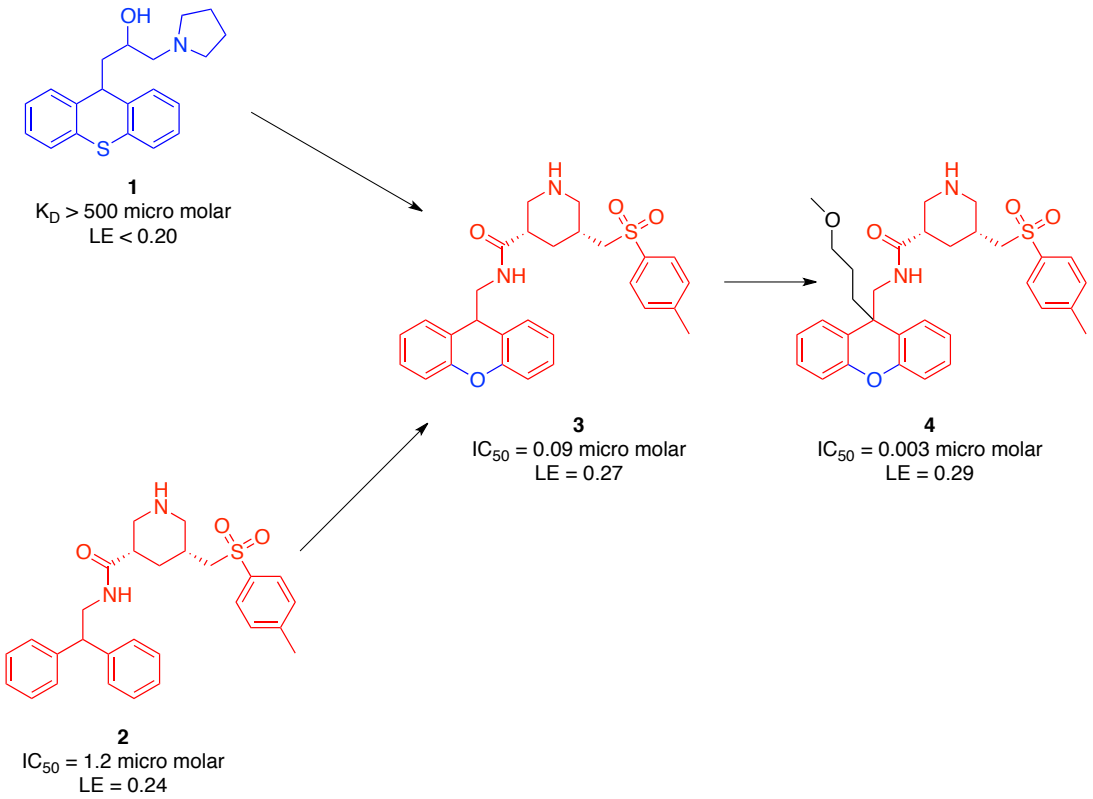
In parallel to our fragment-based lead discovery effort with BioA, the Aldrich laboratory in collaboration with D. Schnappinger has completed screening for functional inhibitors of biotin biosynthesis in *Mtb*[20,73]. In 2011, Dr. Courtney Aldrich's group developed a continuous coupled assay and used this assay to screen a LOPAC library, resulting in the identification of several potent lead inhibitors[73]. Later, the Molecular Libraries and Small Molecules Repository (MLSMR) compound collection of more than 350,000 compounds (PubChem AID#602481) was screened (in press; *Chemistry & Biology*). Also included were 83,000 compounds from the diversity-oriented synthesis (DOS) compound collection. To identify potent inhibitors, they first employed a novel fluorescence-displacement functional HTS assay. This novel assay was a modification of a continuous coupled assay they previously developed to meet the high throughput purpose. From this screening, about 7,600 hit compounds (hit rate 0.095%) with high diversity were identified. To select compounds with cell-based activity that specifically target the biotin biosynthesis pathway, a whole-cell counter-screen was performed in biotin-free and biotin-containing media with isogenic *bioA Mtb* strains that possess differential sensitivity to BioA inhibitors. This counterscreen successfully filtered out off-target hits to allow identification of BioA-specific ligands active in a cellular context. By clustering similar compounds, a number of different inhibitor scaffolds with whole cell activity were identified.

As structural biologists of this team, we determined X-ray crystal structures of BioA complexes with representatives of some of the most potent clusters of compounds from

this screening effort. With the structural analysis, it became clear that the HTS compounds bind in positions that partially overlap fragment binding sites already identified (Chapter 3). This fact allowed us to compare the complex structure of each HTS compound with each fragment complex structure and use fragment merging strategies to optimize the leads from HTS.

In this chapter, we report the structural characterization of six HTS lead compounds in complex with BioA via X-ray crystallography. **H01** through **H05** are HTS lead compounds identified by the *Mtb* whole cell counterscreen; **H08** is one promising lead compound from the miniature LOPAC screening [73]. We then describe our effort to employ structure-based lead optimization strategies that merge details of these complexes with information from fragment hits described in Chapter 3.

Figure 5.1. Optimization of a renin inhibitor by fragment merging.[109]



The fragment (Compound 1, shown in blue) was identified by NMR screening of a focused library. Compound 2 (red) is an HTS hit. Compound 3 was designed using the fragment merging method. The retained features of Compound 2 were shown in red and the modified feature from Compound 1 was shown in blue.

5.2 Experimental procedures

5.2.1. Compound sources.

The seven HTS compounds (**H01**-**H08**) were provided by our collaborator, Dr. Courtney Aldrich. **H01**-**H07** are lead compounds from the MLSMR library; **H08** is from the LOPAC library (Sigma-Aldrich). **H03-1** and **H03-2** were synthesized by Feng Liu in the lab of Prof. Courtney Aldrich.

5.2.2. DSF

The temperature of thermally induced unfolding T_m of high-affinity inhibitors were determined using Differential Scanning Fluorimetry procedures modified from those employed in fragment screening (section 3.2.1.); the ligand concentrations were adjusted because the HTS compounds have much higher affinities than the fragments. 2475 µL of 2X master mix were made at room temperature by mixing 125.0 µL of 1M HEPES pH 7.5, 50.0 µL of 5M NaCl, 5.0 µL of 5000X Sypro orange stock solution, 2295 µL of water. The 2X master mix was cooled on ice. 310 µL of water were distributed into each well of a single column of a 96-well U-bottom plate at room temperature. 19 µL of water were placed into each well of the white PCR plate held by a room-temperature aluminum block. 1 µL of pure DMSO was placed in each well of the first and last column for DMSO-only controls. Using a multi-channel P10 pipette, 1 µL of high affinity ligand (10mg/ml) was placed into a white PCR plate. BioA protein (25 µL of 13.0 A280) was added to the master mix. 310 µL of the 2500 µL 2X master mix was distributed into each well of a single column of a 96-well U-bottom plate on ice. 20 µL of master mix with protein was added into each well of the white PCR plate and gently mixed. Plates were sealed and run in the Bio-Rad CFX96 using 30 sec dwell time per 1°C temperature increase.

5.2.3. Crystallization

BioA was co-crystallized with the HTS leads and the synthesized optimized compounds by the vapor diffusion method in a hanging drop at 20 °C. The crystallization procedure used is as described in Chapter 3 and Chapter 4, except the compound concentration in the initial drops was 0.25 mM.

5.2.4. Data collection, processing and model building

The diffraction data for **H01** and **H08** co-crystals were collected at 100 K using Cu *Kα* radiation on a Rigaku HighFlux HomeLab rotating-anode system with a Saturn 944+ CDD detector in the Kahlert Structural Biology Laboratory at the University of Minnesota. The diffraction data for **H02** and **H03** co-crystals were collected at 100 K with a NOIR-1 CCD at the Advanced Light Source (Berkeley, CA). The diffraction data for **H04**, **H05**, **H03-1** and **H03-2** were collected at 100 K with a Dectris Pilatus 6M Pixel Array Detector on beamline 17-ID (IMCA-CAT) at the Advanced Photon Source, (Argonne, IL). The data for **H01**, **H02**, **H03** and **H08** were processed, integrated, and scaled with d*TREK[111]. The data for **H04**, **H05**, **H03-1** and **H03-2** were processed, integrated, and scaled with XDS[76] and SCALA[88] using the autoPROC scripts available at APS-17-ID (IMCA-CAT). The structures were solved by molecular replacement using Phaser[77] in the CCP4 package[78] using atomic coordinates from PDB code 3TFT as a search model[25]. Refinement and model building was done using Phenix[112] and coot[80]. The figures were prepared with PyMOL (The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC.).

5.2.5. 3D visualization of the binding models of designed compounds

To analyze the binding features of the 6 HTS compounds and compare them with the fragment hits to design next generation compounds, structures were superimposed for

analysis and display using the shared BioA-PLP overlay method of the DrugSite server[81]. Three dimensional models of the designed compounds were built using Prodrug server[110] and Jligand [113]. Bound models of the designed compounds with BioA were built by putting the designed compounds into the electron density generated by their mother template molecules and regularizing the geometry and fit to density in coot.

5.2.6. ITC.

Isothermal Titration Calorimetry was conducted on a GE MicroCal Auto-ITC 200 microcalorimeter. The titration experiment was performed at 25°C in ITC buffer (25 mM Hepes pH 7.5, and 50 mM NaCl). BioA was exchanged into ITC buffer using an Amicon Ultra concentrator, and the final enzyme concentration was determined using the Bradford assay. During the titration, 103 μM compound solution was injected into a solution of the enzyme containing 10.2 μM. The K_A (the association constant in M^{-1}), n (the number of binding sites per monomer) and ΔH (enthalpy) values were determined by ITC. The thermodynamic parameters (ΔG and –TΔS) were calculated using Equation 1.

$$\Delta G = -RTInk = \Delta H - T\Delta S \quad (1)$$

Where ΔG, ΔH and ΔS are the changes in free energy, enthalpy, and entropy of binding, respectively, R = 1.98 cal mol^{-1} K^{-1}, and T is the absolute temperature. The affinities of compounds for BioA are provided as the dissociation constant (K_D =1/K_A). ITC measurements were performed by Dr. Feng Liu in Dr. Aldrich's lab.

5.3 Results

5.3.1. DSF and X-ray crystallography for the HTS lead compounds.

DSF was used to measure the T_m shifts of BioA upon binding with the 7 HTS compounds in Table 5.1. Detailed T_m shifts are listed in Table 5.1. All 7 HTS compounds induced a positive T_m shift equal to or greater than 4 °C. The T_m shifts do not appear to correlate with the IC_{50} data from the continuous fluorescence displacement assay (Table 5.1). Some very potent compounds are tabulated with IC_{50}s below $8x10^{-8}$ M due to the limits in the sensitivity of the HTS assay. A DSF-derived T_m in not available for H08 because this compound was not tested.

We used co-crystallization to try to obtain complex X-ray structures in order to more thoroughly characterize inhibitor binding. Previous experience has shown that co-crystallization results in higher diffraction resolution and higher ligand occupancy. Co-crystallization of all eight compounds was attempted. We were able to obtain complex structures for six out of the eight. Complete electron density is observed for ligands **H01**-**H04**; however for **H05** that is a larger molecule, clear electron density is observed for only part of the molecule (Figure 5.2E). The electron density of the terminal 3-methoxyaniline reflects only partial occupancy in the modeled conformation. The crystallographic statistics for all structures are listed in Table 5.2.

Table 5.1 Structures, DSF T_m shifts and BioA activity from HTS

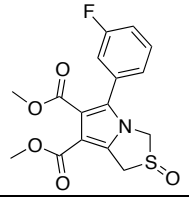
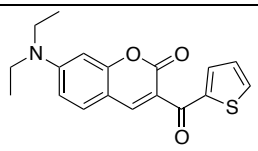
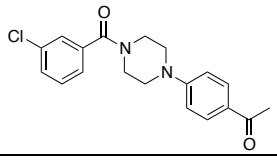
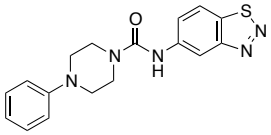
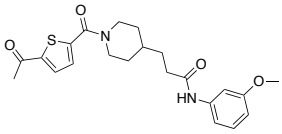
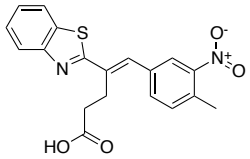
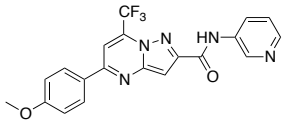
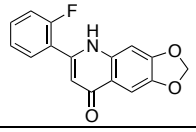
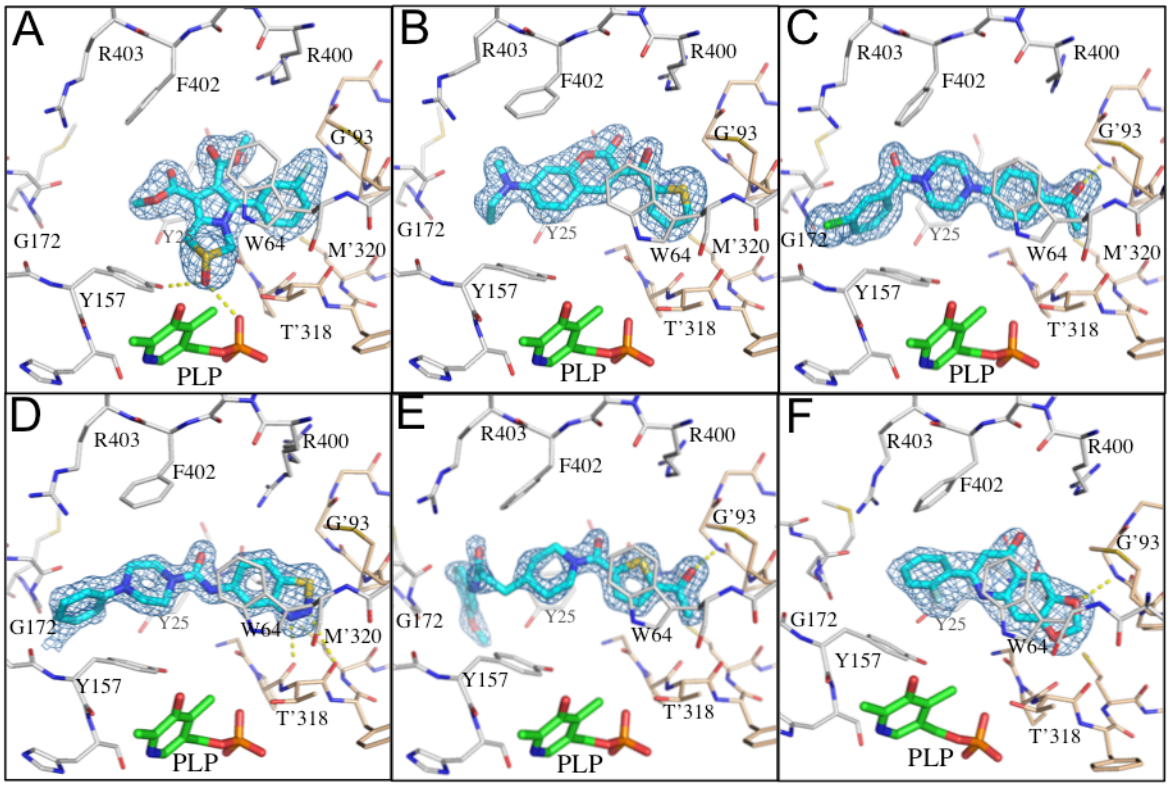
Cmpd ID	Structure	DSF T_m (°C)	ΔT_m (°C)	BioA Activity IC_{50} (M)	X-ray result
H01 BRD-A72996250-001-08-9		93	+8	$1.15x10^{-7}$	Complex
H02 BRD-K50105093-001-06-6		93	+8	$< 8x10^{-8}$	Complex
H03 BRD-K19973577-001-01-3		94	+9	$9.55x10^{-8}$	Complex
H04 BRD-K02445078-001-10-6		94	+9	$< 8x10^{-8}$	Complex
H05 BRD-K57668161-001-08-9		90	+5	$2.25x10^{-7}$	Complex with partial ligand density
H06 BRD-K32354586-001-08-3		91	+6	$< 8x10^{-8}$	
H07 BRD-K93108830-001-01-8		89	+4	$< 8x10^{-8}$	
H08 CHM1				$4.44x10^{-7}$	Complex

Table 5.2 Crystallographic Statistics for Complexes with HTS-derived Compounds

Ligand name	**H01**	**H02**	**H03**	**H04**	**H05**	**H08**
Data collection site	In house	ALS	ALS	IMCAT	IMCAT	In house
Detector	Saturn 944+ CCD	NOIR-1 CCD	NOIR-1 CCD	Dectris Pilatus 6M	Dectris Pilatus 6M	Saturn 944+ CCD
Data set name						
Wavelength (Å)	1.541	1.000	1.000	1.000	1.000	1.541
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions						
***a* (Å)** ***b* (Å)** ***c* (Å)**	63.04 66.24 204.24	63.08 66.21 204.44	63.13 66.04 203.56	63.29 66.01 203.63	63.06 66.05 203.89	62.755, 66.153, 203.261
α,β,γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
molecules per ASU	2	2	2	2	2	2
Resolution (Å)	102.1-2.24 (2.14-2.24)	63.08-1.80 (1.90-1.80)	47.33-1.80 (1.86-1.80)	203.6-1.85 (1.86-1.85)	102.0-1.60 (1.61-1.60)	101.6-2.47 (2.51-2.46)
R_{merge}	0.089 (0.471)	0.076 (0.560)	0.098 (0.529)	0.100 (0.392)	0.058 (0.347)	0.084 (0.083)
I/σ_I	10.2 (1.7)	19.8 (2.6)	9.2 (1.6)	16.0 (6.7)	18.2(4.0)	12.3(6.2)
Completeness	96.1% (81.2%)	90.5% (58.7%)	96.8% (79.0%)	99.8% (100%)	100% (99.9%)	94.08% (79.4%)
Multiplicity	3.20 (1.90)	6.6 (4.2)	6.06 (2.89)	6.1 (6.4)	6.5(5.6)	2.9 (1.7)
No. observations	133962	477296	468895	450465	733585	81740
No. unique reflections	41863	72681	77334	73689	113212	28186
Mosaicity	0.14	0.12	0.4	0.11	0.3	0.5

Refinement	**H01**	**H02**	**H03**	**H04**	**H06**	**H08**
Resolution (Å)	28.19-2.24	102.0-1.90	47.33-1.80	101.82-1.85	101.94-1.60	101.63-2.46
R_{work} **R_{free}**	0.1768 0.2086	0.182 0.221	0.193 0.223	0.191 0.213	0.203 0.220	0.179 0.220
No. atoms	6938	7422	7138	6891	6870	6821
No. water	383	738	481	357	398	212
No. ligand molecule	2	2	2	2	2	2
No. PLP molecule	2	2	2	2	2	2
No. other molecule	6	8	3	2	0	3
Ramachandran plot						
Favored	96.6%	96.6%	96.4%	96.2%	95.9%	94.9%
Allowed	2.7%	2.4%	3.0%	3.2%	3.0%	3.5%
Disallowed	0.7%	1.0%	0.6%	0.6%	1.1%	1.6%
R.m.s deviations						
Bond lengths (Å)	0.003	0.007	0.007	0.008	0.006	0.006
Bond angles (°)	0.95	1.26	1.19	1.21	1.13	1.06

Figure 5.2 HTS Compound Omit Electron Density



Crystal structures for the HTS compounds with 3σ omit density (F_o-F_c) from co-crystal structures of **H01**(A), **H02**(B), **H03**(C), **H04**(D), **H05**(E) and **H08**(F) (cyan), and the molecular environment surrounding the binding site near the PLP (Green). Comparable binding is observed in both active sites of the BioA heterodimer, but only one is shown. Residues from one monomer (colored wheat) are identified with primes; those from the other (colored white) are not.

5.3.2. Structural characterization of the ligand binding

Unambiguous electron density affirms that each inhibitor binds in the hydrophobic site adjacent to the PLP co-factor where substrate KAPA also binds[62](Figure 5.2). This is also the same site into which fragment molecules bind. Inhibitors are in contact with structural components of both monomers, distinguished by color in Figure 5.2. Each inhibitor induces shifts in the conformation of sidechains of Tyr25 and Trp64 to accommodate the longer, flatter molecules which lie sandwiched between repositioned aromatic rings of the tyrosine below and the tryptophan above.

Interactions between the inhibitors and BioA are primarily hydrophobic and arise from two subsites. A subsite on the right (orientation defined by Figure 5.2) is formed by the juxtaposition of Pro24 and Trp 64 with residues 91′-93′ and 316′-318′ (primes denote residues of the other BioA monomer). A subsite on the left exists between the side chain of Tyr25, Tyr157 and the loop formed by Gly172-Met174. A detailed summary of observations regarding the binding of each compound follows.

H01. (Figure 5.2A) The sulfoxide moity of **H01** is in the hydrophobic sub-site composed of Tyr25, Tyr 157, Phe402, Trp64 and Arg403. The core pyrrole is tightly stacked between Tyr25 and Trp64. The Fluorobenzene attached to the pyrrole ring twists a 30^{o} angle with the core and it is interacting with a hydrophobic subsite composed of Met91', Gly93', Gly316' and Arg400. The fluorine is interacting with the hydrophobic Met91, and it can potentially form a fluorine hydrogen bond with the amide backbone of Gly93'. The sulfinyl oxygen is having hydrogen bonding with Tyr157, the phosphate oxygen of PLP and Asp160 bridged with a water molecule. However, the two esters did not make specific interactions. In the **H01** complex, the conformation of the residues in the BioA active site resembles the conformation observed in **F3** and **F9** complexes.

H02. (Figure5.2B) **H02** is composed of two aromatic systems: the core chromenone and the thiophene are linked by a carbonyl group. The thiophene resides in the hydrophobic subsite composed of Met91′, Gly93′, Gly316′ and Arg400, which the fluorobenzene of **H01** interacts with. The core chromenone is stacked between Phe402 and Tyr25. The terminal diethylamine extends into a sub-site composed of Tyr157 and Gly172 backbone. In the **H02** complex, the conformation of the residues in the BioA active site resembles the conformation of the **F3** and **F9** complexes.

H03. (Figure5.2C) **H03** is a linear compound with three hydrophobic rings linked together. It occupies all three sub-sites that **H02** interacts with. The left chloro-benzene ring is in the sub-site the diethylamine of **H02** resides, which is composed of Gly172, Tyr157 and Cys168. The piperazine is in the middle sub-site composed of Tyr25 and Trp64. It has strong hydrophobic interaction with Tyr25. It also serves as a linker between the two aromatic rings. The acetophenyl ring on the right is making π-π interactions with Trp64. The carboxyl oxygen is a hydrogen bond acceptor that interacts with Gly′93 backbone amide. In the **H03** complex, the conformation of the residues in the BioA active site resembles the conformation seen in **F3** and **F9** complexes.

H04. (Figure5.2D) H04 shares similar structural features with **H03**. It is also a linear combination of three hydrophobic rings; the linker moiety is also a piperazine with a carbonyl group. However, **H04** is not binding in the same fashion as **H03**; it is 180^{o} flipped from **H03**. There are two major differences between **H03** and **H04**: the right hand side acetophenyl ring of **H03** was replaced by a phenyl ring in **H04**, and the left hand side chlorophenyl was replaced by a benzothiadiazole. This flip enabled **H04** to have

more favorable interactions, because the right hand side sub-site(which is composed of Met91′, Gly93′, Gly316′ and Arg400) has a hydrogen bond donor(backbone amide of Gly93′), and the benzothiadiazole can have hydrogen bonds with Gly93′. In addition, the flipped binding conformation of **H04** allowed its benzothiadiazole to accommodate a better Pi-Pi interaction with Trp64. One the other hand, the phenyl ring also extended deeply into the left sub-site which is composed of Tyr157, Gly172' and Gly156. In the **H04** complex, the conformation of the residues in the BioA active site resembles the conformation seen in the **F2** complex.

H05. (Figure5.2E) H05 is composed of three hydrophobic ring systems. It has a four-atom linker between the piperidine and the terminal phenyl ring, which adds more flexibility to this compound. The thiophene and the piperidine are interacting with the right hand side hydrophobic subsite of BioA in the similar fashion as **H03** and **H04**: the right hand side carbonyl oxygen has a hydrogen bond with Gly′93 backbone amide; the thiophene resides in the sub-site composed of Trp64 and Pro24, having π-π interactions with Trp64; the piperidine goes through a wide tunnel composed of Tyr157, Phe402 and Tyr25. The terminal methoxyphenyl ring is linked with the piperidine by a C4 linker. The density of the methoxyphenyl is not very complete. From the density model, we can see that, the terminal methoxyphenyl exists in two different conformations and potentially targets two different sub-sites. In one conformation, the methoxyphenyl resides in the sub-site composed of Gly172, Tyr157 and Cys168, which **H03** and **H04** also interact with. In another conformation, which is the major conformation, **H05** methoxyphenyl extended into a sub-site composed of Tyr25, Met174 and Trp178. The methoxy oxygen makes a hydrogen bond with the indole of Trp178. In the **H05** complex, the conformation

of the residues in the BioA active site resembles the conformation observed in the **F3** complex.

H08. (Figure5.2F) H08 has very strong hydrophobic interactions with the hydrophobic sub-site composed of Trp64, Pro24, Tyr25 and Phe402. Trp64 and Tyr25 side chains make π-π interactions with the quinoline moiety and the fluorophenyl moiety of **H08**. The Tyr25 side chain is shifted about 1 Å towards the fluorophenyl ring and the phenyl plane is rotated through a 30^0 angle to accommodate the π-π interaction. The fluorine induces a 15^o rotation angle between the two ring systems, which allows both rings to have π-π interaction with the surrounding residues. **H08** also has a hydrogen bond with Gly′93 backbone amide. In the **H05** complex, the conformation of the residues in the BioA active site resembles the conformation seen in the **F2** complex.

5.3.3. Optimization of the HTS compounds based on fragment merging-molecular design

As the HTS compounds bind in the same subsite as the fragments bind, we were able to overlay the fragment structures and HTS compound structures to seek opportunities to combine features of two molecules to gain affinity and selectivity. However, as we overlay the structures one by one, we discovered that it is not easy to find components of different molecules that align well. This is not surprising. Protein conformational flexibility aggravates the unalignable vector problem, because the binding site actually drifts as large hydrophobic side chains in the binding site (Trp 64, Trp 65, Tyr 25, Phe402) change conformation. Despite of these facts, we were able to identify some possibilities for fragment merging.

Optimization opportunities for H01:

Following an overlay of the **H01** complex structure with the complex with **F3**, the hydrophobic cores of the two molecules align in the same spot; the angles between the two hydrophobic rings in the two molecules are also the same (Figure 5.3A). This comparison led to the suggestion that the core aromatic ring scaffold can be changed, Derivatives of **H01** with altered aromatic rings in the hydrophobic core were designed (Figure 5.3A **H01.1**-**H01.5**).

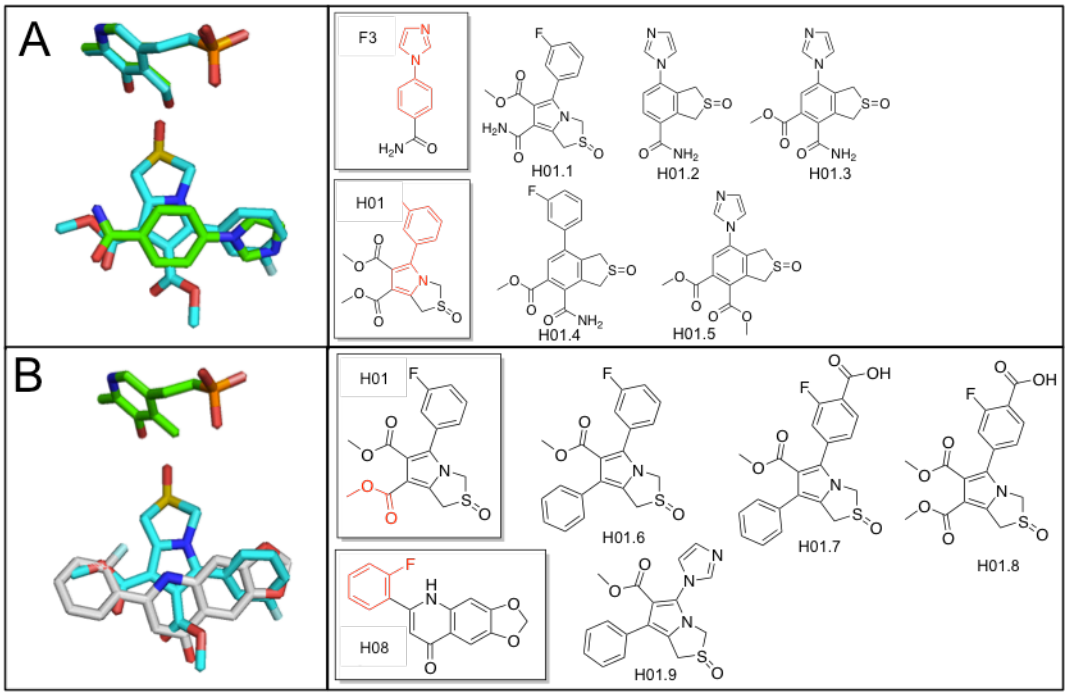
In overlaying complexes of **H01** and **H08**, we observed that, although both compounds have a fluorobenzene ring, they do not overlap; instead, the **H08** fluorobenzene ring overlaps with an ester in **H01** (Figure 5.3B). Substitution of this **H01** ester with an aromatic ring could potentially increase binding affinity by improving the π-π interaction with Y25 (Figure 5.3B **H01.6**, **H01.7**, **H01.9**). From modeling, we predict that the phenyl ring will adopt a good orientation for effective π-π stacking against Tyr25 (Figure 5.4B). The fluorobenzene of **H01** interacts with the right hand side hydrophobic site in BioA. This is the same sub-site occupied by the dioxole in the complex with **H08**. There is extra space in this sub-site that should permit the addition of a polar substituent (i.e., a carboxylate) to the para position of the benzene ring (Figure 5.3B **H01.8**, Figure 5.4C). Such a carboxylate should be positioned upon binding to form a specific hydrogen bond with the Gly′93 backbone amide. From modeling, the distance between the carbonyl oxygen and the Gly′93 backbone nitrogen is 3.1Å (Figure 5.4C), and the C-O-N angle was 81.1^0. The angle is not very favorable for the hydrogen bond; the flexibility of the binding site to accommodate this compound needs to be further tested.

Optimization opportunities for H03:

Figure 5.5A illustrates the excellent coincidence of atoms in the complexes of HTS compound **H03** and **F9**. The overlap is good not only on the aromatic core, but also in the coincidence of polar atoms. Grafting a five-membered ring onto the **H03** acetylphenone would potentially strengthen the π-π interaction relating this aromatic ring system to the Trp64 indole, as these two rings are already parallel with each other (figure 5.5A-C). In addition, the carbon on the meta-position of the fluorobenzene ring is 3.2 Å away from the hydroxyl oxygen of Tyr25 (Figure 5.5D). The Tyr25 hydroxyl can either be a hydrogen bond donor or acceptor. This carbon could be replaced with a heteroatom such as nitrogen or oxygen to gain an additional hydrogen bond (Figure 5.5D).

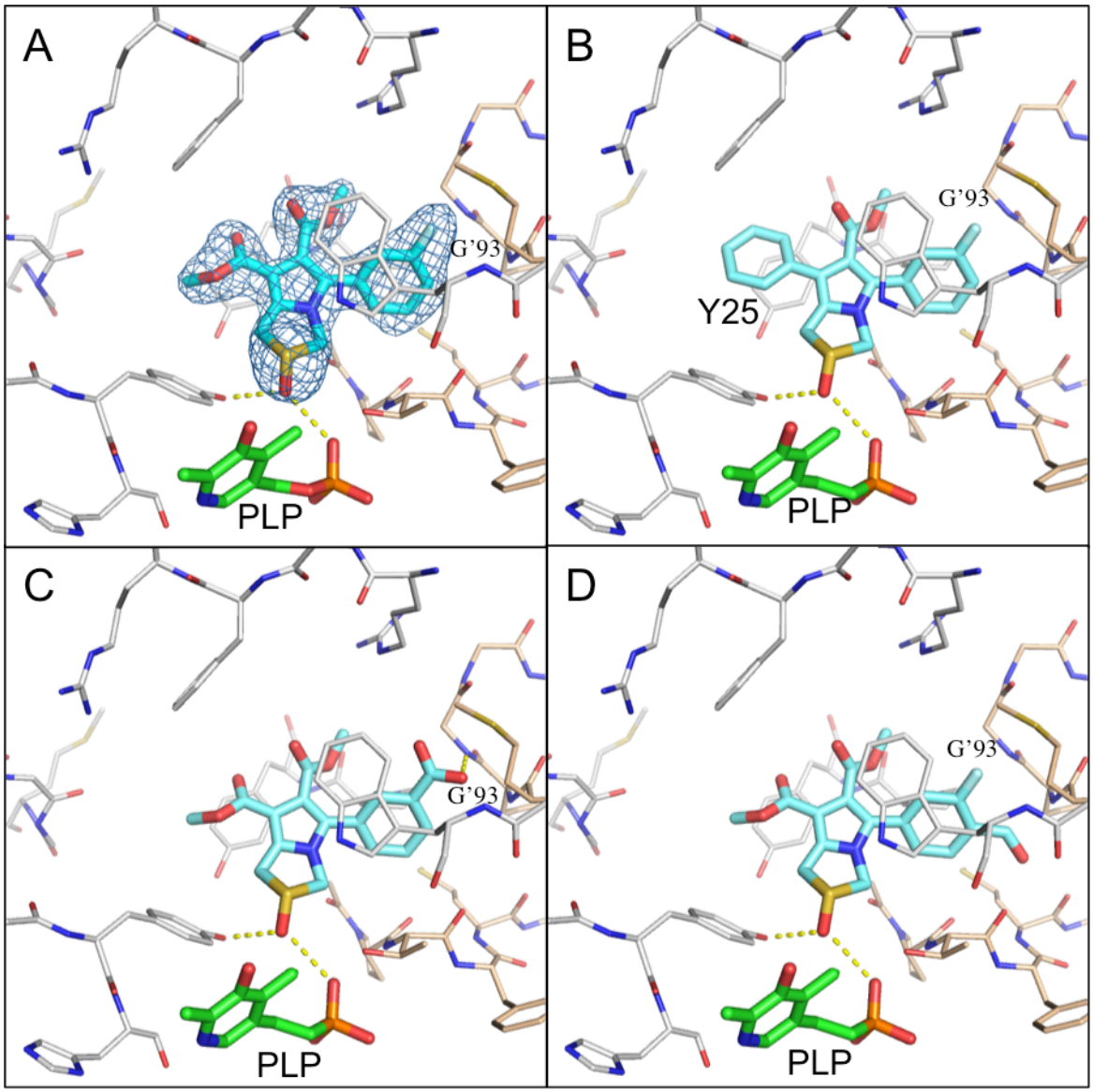
Compounds designed based on our structure-based and fragment-based approaches are being synthesized in the Aldrich laboratory. When synthesis is completed, analogs will be characterize by X-ray crystallography and bioassays and compared to design expectations.

Figure 5.3 BioA inhibitor design by merging H01 with F3 or H08.



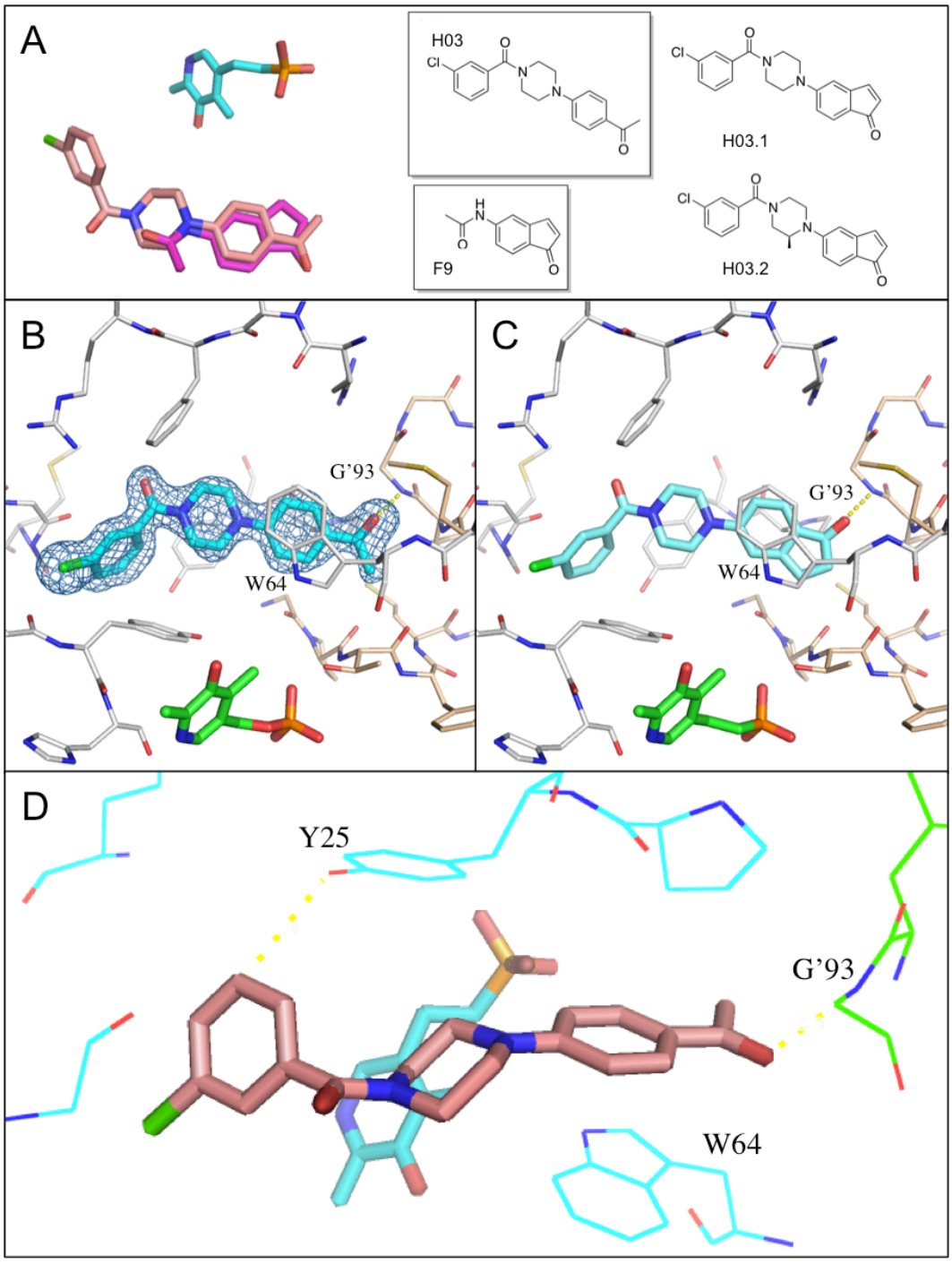
A) **H01** and **F3** overlaid in the BioA active site, and corresponding designed inhibitor examples; B) **H01** and **H08** overlaid in the BioA active site, and corresponding designed inhibitor examples.

Figure 5.4 Binding models of the designed compounds made using coot.



A) BioA-**H01** X-ray crystal structure; B) model of inhibitor designed from Merging features of **H01** and **H08** (**H01.6**); C) and D) model of inhibitor designed from **H01** to gain an extra hydrogen bond with Gly'93 backbone nitrogen (**H01.8**).

Figure 5.5 BioA inhibitor designs optimizing H03.



A) Overlaid **H03** and **F9** complexes, and designed inhibitor conceptions; B) BioA-**H01** X-ray crystal structure; C) Model of compound designed by merging **H03** with **F9**; D) Rationale for changing the meta carbon on the fluorobenzene to a heteroatom.

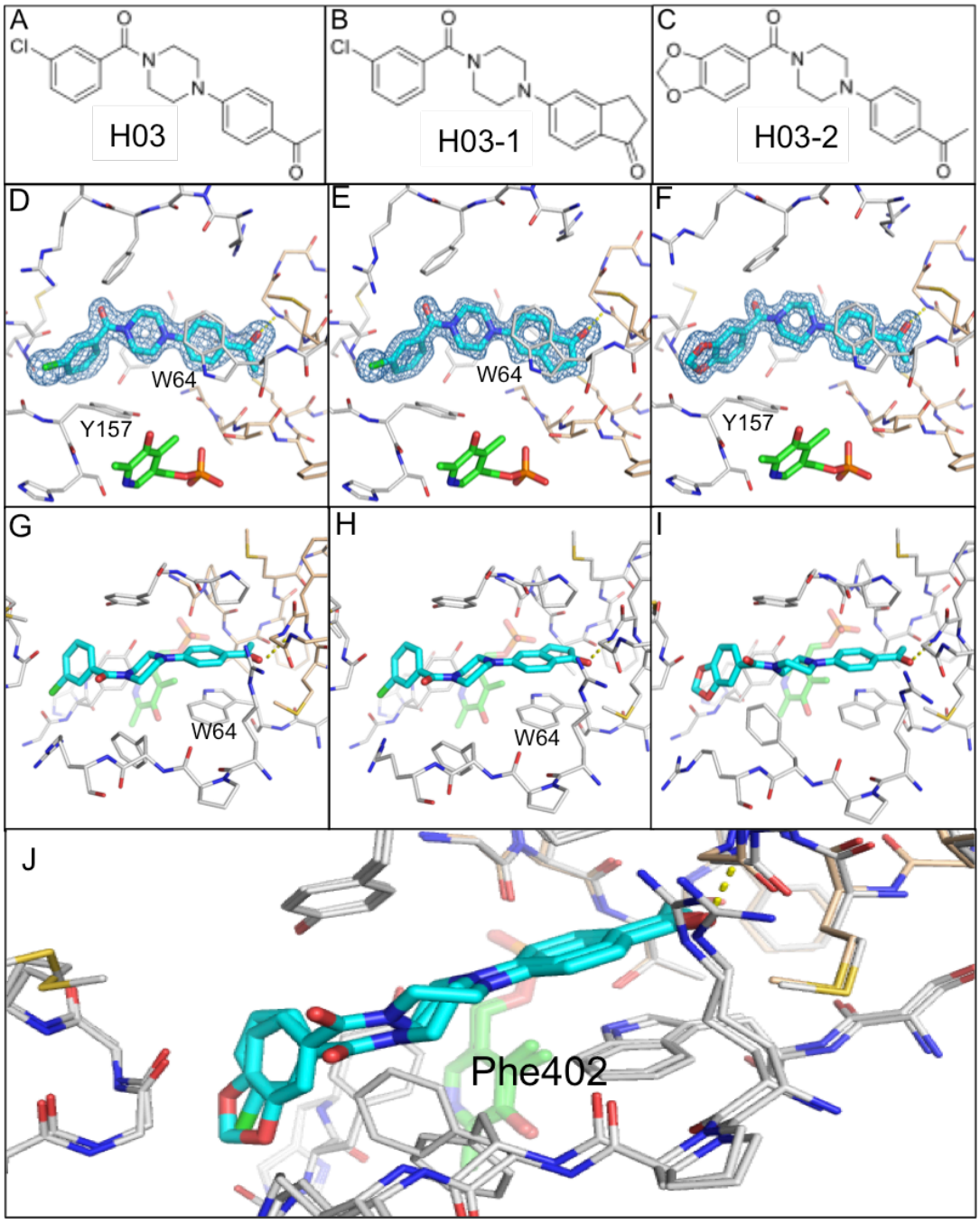
5.3.4. Synthesis and characterization of the optimized compounds from HTS.

Independently, collaborators in Aldrich lab have performed more general scaffold hopping, and made a variety of analogs of the most potent HTS leads. Compounds **H03-1** (Figure 5.6B) and **H03-2** (Figure 5.6C) are two of such analogs synthesized by Feng Liu in the Aldrich lab. These two compounds have been characterized by X-ray crystallography (Figure 5.6E and F) and ITC experiments. **H03-1** and **H03-2** structures were both obtained by co-crystallization. Co-crystals with **H03-1** and **H03-2** diffracted to resolutions of 1.5 Å and 1.6 Å respectively; X-Ray data collection and refinement statistics show that these structures are well-refined (Table 5.3).

The bound conformation of **H03-1** is the same as that found with the original HTS template **H03**. The cyclopentanone ring added in place of the acetylphenyl group makes an extended hydrophobic interaction with the aromatic Trp64 side chain as predicted (Figure 5.6E, H). Encouragingly, this crystallographic structure closely resembles the predicted binding conformation for one of our designed cyclopentadienone inhibitors (Figure 5.5C). From the ITC experiment, we observed a 3-fold decrease of K_D value (654 nM for **H03** *vs*. 211 nM for **H03-1**) (Table 5.4).

H03-2 was designed by using a benzodioxole ring to take place of the fluorobenzene ring of **H03** (Figure 5.6C). From the X-ray structure, we can see that the orientation of the molecule in the active site remained the same as **H03**, however, the conformation of the piperazine ring changed from a chair to a twist (Figure 5.6G, I, J). The oxygen on the carbonyl group which links the piperazine ring and the fluorobenzene ring shifted towards Phe402 by 1.7 Å and the benzene ring also rotated 15^o due to the conformation change. Phe402 Phenyl ring rotated 45^0 to make room for these changes.

Figure 5.6 X-ray crystallography structures of H03, H03-1 and H03-2.



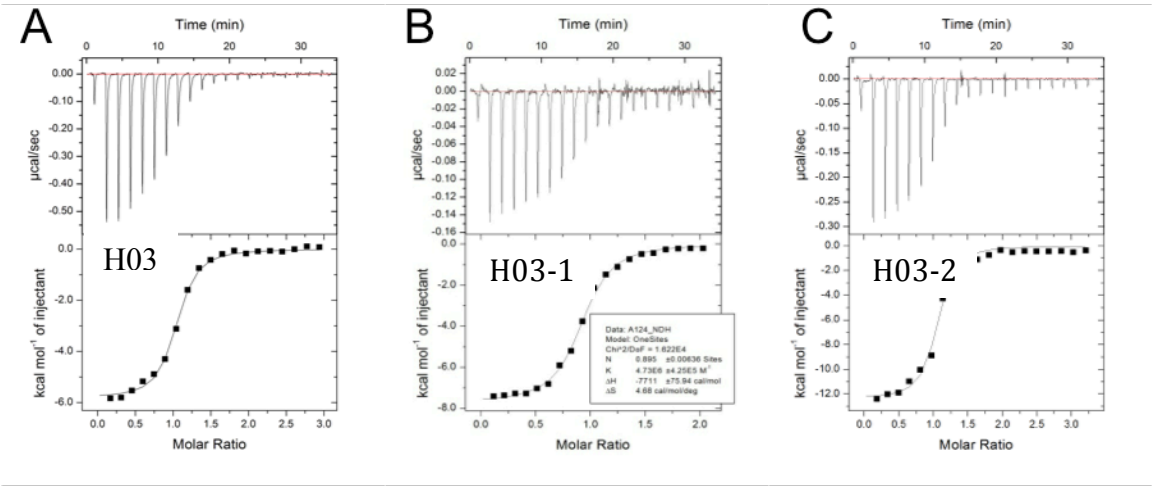
A-C) Structures of **H03**, **H03-1** and **H03-2**; D-F) co-crystal structures of **H03**, **H03-1** and **H03-2** with 3σ omit density (F_o-F_c) maps; G-I) co-crystal structures of **H03**, **H03-1** and **H03-2** in a different view; J) An overlaid view of **H03** (G) and **H03-2** (I) in BioA active site.

Table 5.3 Crystallographic statistics for complex structures of H03-1 and H03-2

Ligand name	**H03-1**	**H03-2**
Data collection site	IMCAT	IMCAT
Detector	Dectris Pilatus 6M	Dectris Pilatus 6M
Wavelength (Å)	1.000	1.000
Space group	P212121	P212121
Cell dimensions		
a, b, c (Å)	63.23 66.03 203.26	63.47 65.92 204.25
α,β,γ (°)	90 90 90	90 90 90
molecules per ASU	2	2
Resolution (Å)	101.63-1.50 (1.51-1.50)	204.25-1.60 (1.61-1.60)
R_{merge}	0.082(0.319)	0.097(0.411)
I/σ_I	14.0(5.0)	13.3(4.9)
Completeness	99.5% (96.6%)	99.7% (99.7%)
Multiplicity	6.4(6.7)	6.4(6.5)
No. observations	878235 (8862)	729227 (7586)
No. unique reflections	136219 (1325)	113574 (1175)
Refinement		
Resolution (Å)	33.87-1.50	32.54-1.60
R_{work}/R_{free}	0.1680/0.1913	0.1791/0.2014
No. atoms	7317	6944
No. water	654	328
No. ligand molecule	2	2
No. PLP molecule	2	2
No. other molecule	0	0
Ramachandran plot		
Favored	97.4%	97.7%
Allowed	2.0%	1.7%
Disallowed	0.6%	0.6%
R.m.s deviations from ideal geometry		
Bond lengths (Å)	0.007	0.007
Bond angles (°)	1.24	1.21

Table 5.4 ITC data for H03 (A), H03-1 (B) and H03-2 (C).

	N	$K_{D,}$ nM	ΔH, kcal mol^{-1}	ΔG, kcal mol^{-1}	-TΔS, kcal mol^{-1}
H03	0.998	654	-5.8	- 8.4	2.6
H03-1	1.00	211	-7.71	- 9.07	1.36
H03-2	1.00	103	-12.37	- 9.49	2.88



5.4 Discussion

5.4.1. BioA co-crystallization with ligands

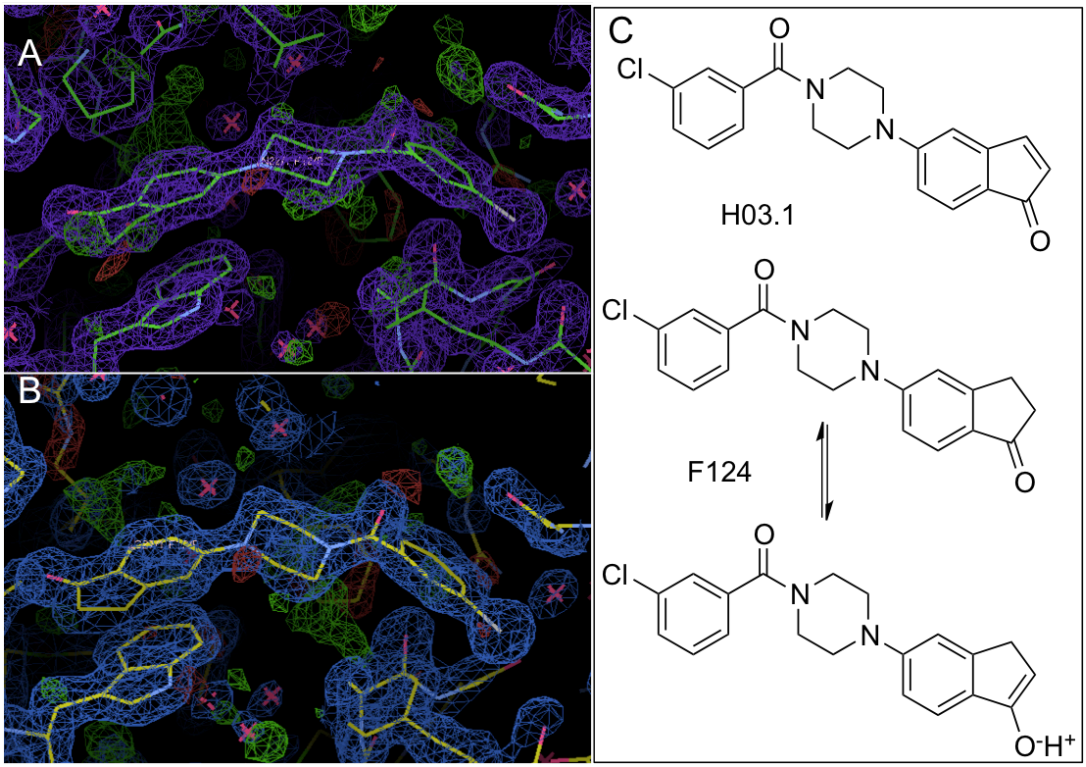
All structures described herein involving higher-potency inhibitors of BioA were obtained from crystals derived from co-crystallization rather than soaking. Soaking protocols successfully employed in the preparation of fragment-bound crystals are described in Chapter 3. Both approaches were used to prepare crystals with more potent compounds, but co-crystallization produced better results. Diffraction data obtained from holo BioA crystals soaked with **H02**-**H05** gave rise to no observable ligand electron density, and electron density quality for **H01** was not as good when obtained using soaked crystals. The preparation of BioA crystals by co-crystallization is slightly slower than the preparation of holo BioA crystals; co-crystals appear in 72 hours and grow to their full size in 7 days while BioA holo crystals grow to their full size in 3 days. The quality of the crystals is improved when co-crystallized with the HTS compounds, so the investment of additional time and protein is justified.

5.4.2. Conformation of the cyclopentanone in **H03-1**

A cyclopentanone is fused on the benzene group in compound **H03-1**. A cyclopentanone is usually stable in an “envelope” or half-chair conformation. However, in the complex with bound **H03-1**, this 5-membered ring is in a planar conformation. We confirmed this planar conformation using the electron density generated from our X-ray data. We have modeled this cyclopentanone in the two different conformations to see which one best fit the electron density. When we model the density using an envelope shaped conformation of the cyclopentanone, we can see unmodeled density for the 5-membered ring (Figure 5.7A shown in green). When modeled as a planar 5-membered ring, the structure fits the density perfectly (Figure 5.7B). This cyclopentanone is capable of a

keto-enol tautomerization (Figure 5.7C). This tautomerization alters the functional group hydrogen bonding capability. The keto form can only be an acceptor, while the enol form can potentially be a donor or an acceptor.

Figure 5.7 Conformation of the cyclopentanone in H03-1.



The **H03-1** cyclopentanone is better fit into the planar conformation model. A) An envelope conformation modeled to fit the ligand density; B) A planar conformation modeled to fit the ligand density; C) Upper: inhibitor (**H03.1**) designed by merging **H03** with **F9**. Lower: the ketone-enol tautomerization of the cyclopentanone in **H03-1**. Since the cyclopentanone is in a planar conformation, its binding structure information can give us a clue of how **H03.1** would bind.

5.4.3. Improved inhibitor binding affinity by optimization based on fragment merging

H03-1 is a very nice example of fragment-based drug discovery, in which fragment information is used as one aspect of an overall lead discovery program. However, such opportunities are not always easy to find. More often it is hard to find common binding features of two hit molecules, also there are cases after merging the features, the designed compound does not bind in the predicted conformation. From our design experience, more iterations of modification with smaller modification in each step can increase the success rate in the optimization process. As we have obtained improved BioA inhibitors using fragment merging (**H03-1** and **H03-2**), we can continue to merge features of **H03-1** and **H03-2** to aim for further increase in activity.

5.5 Conclusion

In this chapter, we described our biophysical and structural characterization of some BioA inhibitors from HTS. The DSF results showed that all of the HTS hits are stabilizing BioA protein in solution. Further validation using X-ray crystallography confirmed that 6 out of the 8 HTS compounds bind in the BioA active site. We used our structure analysis and fragment merging strategy to propose optimized inhibitors for synthesis. Two analogs of HTS hit **H03**, one of them based on our design (**H03-1**), were synthesized by our collaborator Dr. Feng Liu. We characterized the optimized compounds using X-ray crystallography and ITC. The observed binding is very much as predicted based on modeling. We also observed increased binding affinity as determined by calorimetric measurements.

Chapter VI: Conclusion

The fragment based inhibitor design project starts with a three-stage fragment screening. Initial DSF screening reduced our workload for structural characterization by identifying 21 initial hits from the 1000 compound library. From the complex structures from X-ray crystallography we obtained the most important information for further molecular design. ITC was then used to analyze the thermodynamics of compound binding. Data from ITC not only enable us to rank the hit compounds by binding affinity, but more importantly, it can be correlated with the structural information to understand how different interactions and different conformational changes affect the enthalpy and entropy of the target.

Although strategies for FBDD are most commonly categorized into 3 classes (fragment extension, fragment linking and fragment merging), the use of the fragment information should never be limited to these classified methods. Creative and useful methods can be developed when two concepts are combined together. The fragment hits all bind in the same sub-site of BioA, which makes fragment linking very hard. Thus we focused on small modification and extension of the fragments in Chapter 4, and discovered a novel mechanism based inhibitor **F10-1** and an analog of **F5** with 2 fold increased binding affinity (**F5-1**). In chapter 5, we showed examples of merging fragments with HTS compounds to perform rational scaffold hopping. HTS analog **H03-1** is very similar in structure to several of those we have designed via fragment merging, and our structural and biophysical characterization gives us confidence that our designed inhibitor **H03.1** will bind as predicted.

The completion of this thesis does not represent a completion of this project, but rather a beginning. Incremental modifications of known BioA inhibitors should continue in small steps, considering the flexibility of the BioA active site. Most of the compounds examined so far have been available from commercial sources, but structural data generated here establishes considerable opportunities for the design and synthesis of novel and specific BioA inhibitors with better therapeutic potential.

References

1. Dye, C.; Lonnroth, K.; Jaramillo, E.; Williams, B. G.; Raviglione, M. Trends in Tuberculosis Incidence and Their Deter Determinants in 134 Countries. In *Bull World Health Organ*; Vol. 87, pp. 683–691.

(2) Cooper, A. M. Cell-Mediated Immune Responses in Tuberculosis. *Annu. Rev. Immunol.* **2009**, *27*, 393–422.

(3) Barry, C. E.; Boshoff, H. I.; Dartois, V.; Dick, T.; Ehrt, S.; Flynn, J.; Schnappinger, D.; Wilkinson, R. J.; Young, D. The Spectrum of Latent Tuberculosis: Rethinking the Biology and Intervention Strategies. *Nat Rev Micro* **2009**, *7*, 845–855.

(4) Young, D.; Dye, C. The Development and Impact of Tuberculosis Vaccines. *Cell* **2006**, *124*, 683–687.

(5) Iseman, M. D. Tuberculosis Therapy: Past, Present and Future. *Eur. Respir. J.* **2002**, *20*, 87S – 94s.

(6) Chakraborty, S.; Gruber, T.; Barry, C. E.; Boshoff, H. I.; Rhee, K. Y. Para-Aminosalicylic Acid Acts as an Alternative Substrate of Folate Metabolism in Mycobacterium Tuberculosis. *Science* **2013**, *339*, 88–91.

(7) Zheng, J.; Rubin, E. J.; Bifani, P.; Mathys, V.; Lim, V.; Au, M.; Jang, J.; Nam, J.; Dick, T.; Walker, J. R.; Pethe, K.; Camacho, L. R. Para-Aminosalicylic Acid Is a Prodrug Targeting Dihydrofolate Reductase in Mycobacterium Tuberculosis. *J. Biol. Chem.* **2013**, *288*, 23447–23456.

(8) Mitchison, D. A. The Diagnosis and Therapy of Tuberculosis During the Past 100 Years. *Am. J. Respir. Crit. Care Med.* **2005**, *171*, 699–706.

(9) Controlled Clinical Trial of Four Short-Course (6-Month) Regimens of Chemotherapy for Treatment of Pulmonary Tuberculosis. In; 1974; pp. 1100–1106.

(10) Forbes, M.; Kuck, N.; Peets, E. Mode of Action of Ethambutol. *J Bacteriol* **1962**, *84*, 1099–1103.

(11) Zimhony, O.; Cox, J. S.; Welch, J. T.; Vilcheze, C.; Jacobs, W. R. Pyrazinamide Inhibits the Eukaryotic-like Fatty Acid Synthetase I (FASI) of Mycobacterium Tuberculosis. *Nat Med* **2000**, *6*, 1043–1047.

(12) Caminero, J. A.; Sotgiu, G.; Zumla, A.; Migliori, G. B. Best Drug Treatment for Multidrug-Resistant and Extensively Drug-Resistant Tuberculosis. *Lancet Infect. Dis.* **2010**, *10*, 621–629.

(13) Kwon, Y. S.; Kim, Y. H.; Suh, G. Y.; Chung, M. P.; Kim, H.; Kwon, O. J.; Choi, Y. S.; Kim, K.; Kim, J.; Shim, Y. M.; Koh, W.-J. Treatment Outcomes for HIV-Uninfected Patients with Multidrug-Resistant and Extensively Drug-Resistant Tuberculosis. *Clin. Infect. Dis.* **2008**, *47*, 496–502.

(14) Menzies, D.; Benedetti, A.; Paydar, A.; Martin, I.; Royce, S.; Pai, M.; Vernon, A.; Lienhardt, C.; Burman, W. Effect of Duration and Intermittency of Rifampin on Tuberculosis Treatment Outcomes: A Systematic Review and Meta-Analysis. *PLoS Med.* **2009**, *6*, e1000146.

(15) *Treatment of Tuberculosis. Guidlines for National Programmes, 2nd Ed.*; WHO/TB/97.220; Stop TB, World Health Organization, 1997.

(16) Zhou, S.-F.; Xue, C.; Yu, X.-Q.; Li, C.; Wang, G. Clinically Important Drug Interactions Potentially Involving Mechanism-Based Inhibition of Cytochrome P450 3A4 and the Role of Therapeutic Drug Monitoring. *Drug Monit* **2007**, *29*, 687–710.

(17) Davies, G.; Cerri, S.; Richeldi, L. Rifabutin for Treating Pulmonary Tuberculosis. *Cochrane Database Syst. Rev.* **2007**, CD005159.

(18) Ioerger, T. R.; O'Malley, T.; Liao, R.; Guinn, K. M.; Hickey, M. J.; Mohaideen, N.; Murphy, K. C.; Boshoff, H. I. M.; Mizrahi, V.; Rubin, E. J.; Sassetti, C. M.; Barry III, C. E.; Sherman, D. R.; Sacchettini, J. C. Identification of New Drug Targets and Resistance Mechanisms in Mycobacterium Tuberculosis. *PLoS ONE* **2013**, *8*, e75245.

(19) Payne, D. J.; Gwynn, M. N.; Holmes, D. J.; Pompliano, D. L. Drugs for Bad Bugs: Confronting the Challenges of Antibacterial Discovery. *Nat Rev Drug Discov* **2007**, *6*, 29–40.

(20) Sassetti, C. M.; Rubin, E. J. Genetic Requirements for Mycobacterial Survival during Infection. *Proc. Natl. Acad. Sci.* **2003**, *100*, 12989–12994.

(21) Dey, S.; Lane, J. M.; Lee, R. E.; Rubin, E. J.; Sacchettini, J. C. Structural Characterization of the Mycobacterium Tuberculosis Biotin Biosynthesis Enzymes 7,8-Diaminopelargonic Acid Synthase and Dethiobiotin Synthetase,. *Biochemistry (Mosc.)* **2010**, *49*, 6746–6760.

(22) Sae Woong Park; Klotzsche, M.; Wilson, D. J.; Boshoff, H. I.; Hyungjin Eoh; Manjunatha, U.; Blumenthal, A.; Kyu Rhee; Barry III, C. E.; Aldrich, C. C.; Ehrt, S.; Schnappinger, D. Evaluating the Sensitivity of Mycobacterium Tuberculosis to Biotin Deprivation Using Regulated Gene Expression. *PLoS Pathog.* **2011**, *7*, 1–10.

(23) Mann, S.; Ploux, O. 7,8-Diaminoperlargonic Acid Aminotransferase from Mycobacterium Tuberculosis, a Potential Therapeutic Target. *FEBS J.* **2006**, *273*, 4778–4789.

(24) Eisenberg, M.; Stoner, G. Biosynthesis of 7,8-Diaminopelargonic Acid, a Biotin Intermediate, from 7-Keto-8-Aminopelargonic Acid and S-Adenosyl-L-Methionine. *J. Bacteriol.* **1971**, *108*, 1135–1140.

(25) Shi, C.; Geders, T. W.; Park, S. W.; Wilson, D. J.; Boshoff, H. I.; Abayomi, O.; Barry, C. E.; Schnappinger, D.; Finzel, B. C.; Aldrich, C. C. Mechanism-Based Inactivation by Aromatization of the Transaminase BioA Involved in Biotin Biosynthesis in Mycobaterium Tuberculosis. *J. Am. Chem. Soc.* **2011**, *133*, 18194–18201.

(26) Käck, H.; Sandmark, J.; Gibson, K.; Schneider, G.; Lindqvist, Y. Crystal Structure of Diaminopelargonic Acid Synthase: Evolutionary Relationships between Pyridoxal-5′-Phosphate-Dependent Enzymes. *J. Mol. Biol.* **1999**, *291*, 857–876.

(27) Hotta, K.; Kitahara, T.; Okami, Y. Studies of the Mode of Action of Amiclenomycin. *J. Antibiot. (Tokyo)* **1975**, *28*, 222–228.

(28) Sandmark, J.; Mann, S.; Marquet, A.; Schneider, G. Structural Basis for the Inhibition of the Biosynthesis of Biotin by the Antibiotic Amiclenomycin. *J. Biol. Chem.* **2002**, *277*, 43352–43358.

(29) Shi, C.; Aldrich, C. C. Design and Synthesis of Potential Mechanism-Based Inhibitors of the Aminotransferase BioA Involved in Biotin Biosynthesis. *J. Org. Chem.* **2012**, *77*, 6051–6058.

(30) Keseru, G. M.; Makara, G. M. The Influence of Lead Discovery Strategies on the Properties of Drug Candidates. *Nat Rev Drug Discov* **2009**, *8*, 203–212.

(31) Erlanson, D. Introduction to Fragment-Based Drug Discovery. In *Fragment-Based Drug Discovery and X-Ray Crystallography*; Davies, T. G.; Hyvönen, M., Eds.; Topics in Current Chemistry; Springer Berlin Heidelberg, 2012; Vol. 317, pp. 1–32.

(32) Nasiri, H. R.; Bell, N. M.; McLuckie, K. I. E.; Husby, J.; Abell, C.; Neidle, S.; Balasubramanian, S. Targeting a c-MYC G-Quadruplex DNA with a Fragment Library. *Chem. Commun.* **2014**, *50*, 1704–1707.

(33) Abad-Zapatero, C.; Metz, J. T. Ligand Efficiency Indices as Guideposts for Drug Discovery. *Drug Discov. Today* **2005**, *10*, 464–469.

(34) Hopkins, A. L.; Groom, C. R.; Alex, A. Ligand Efficiency: A Useful Metric for Lead Selection. *Drug Discov. Today* **2004**, *9*, 430–431.

(35) Rees, D. C.; Congreve, Miles; Murray, C. W.; Carr, R. Fragment-Based Lead Discovery. *Nat Rev Drug Discov* **2004**, *3*, 660–672.

(36) Shuker, S. B.; Hajduk, P. J.; Meadows, R. P.; Fesik, S. W. Discovering High-Affinity Ligands for Proteins: SAR by NMR. *Science* **1996**, *274*, 1531–1534.

(37) Wyss, D. F.; Arasappan, A.; Senior, M. M.; Wang, Y.-S.; Beyer, B. M.; Njoroge, F. G.; McCoy, M. A. Non-Peptidic Small-Molecule Inhibitors of the Single-Chain Hepatitis C Virus NS3 Protease/NS4A Cofactor Complex Discovered by Structure-Based NMR Screening. *J. Med. Chem.* **2004**, *47*, 2486–2498.

(38) Wu, B.; Zhang, Z.; Noberini, R.; Barile, E.; Giulianotti, M.; Pinilla, C.; Houghten, R. A.; Pasquale, E. B.; Pellecchia, M. HTS by NMR of Combinatorial Libraries: A Fragment-Based Approach to Ligand Discovery. *Chem. Biol.* **2013**, *20*, 19–33.

(39) Whiting, M.; Muldoon, J.; Lin, Y.-C.; Silverman, S. M.; Lindstrom, W.; Olson, A. J.; Kolb, H. C.; Finn, M. G.; Sharpless, K. B.; Elder, J. H.; Fokin, V. V. Inhibitors of HIV-1 Protease by Using In Situ Click Chemistry. *Angew. Chem.* **2006**, *118*, 1463–1467.

(40) Howard, N.; Abell, C.; Blakemore, W.; Chessari, G.; Congreve, M.; Howard, S.; Jhoti, H.; Murray, C. W.; Seavers, L. C. A.; van Montfort, R. L. M. Application of Fragment Screening and Fragment Linking to the Discovery of Novel Thrombin Inhibitors†. *J. Med. Chem.* **2006**, *49*, 1346–1355.

(41) Greenblatt, H. M.; Guillou, C.; Guénard, D.; Argaman, A.; Botti, S.; Badet, B.; Thal, C.; Silman, I.; Sussman, J. L. The Complex of a Bivalent Derivative of Galanthamine with Torpedo Acetylcholinesterase Displays Drastic Deformation of the Active-Site Gorge: Implications for Structure-Based Drug Design. *J. Am. Chem. Soc.* **2004**, *126*, 15405–15411.

(42) Beevers, R. E.; Buckley, G. M.; Davies, N.; Fraser, J. L.; Galvin, F. C.; Hannah, D. R.; Haughan, A. F.; Jenkins, K.; Mack, S. R.; Pitt, W. R.; Ratcliffe, A. J.; Richard, M. D.; Sabin, V.; Sharpe, A.; Williams, S. C. Low Molecular Weight Indole Fragments as IMPDH Inhibitors. *David W Robertson Meml. Issue* **2006**, *16*, 2535–2538.

(43) Beevers, R. E.; Buckley, G. M.; Davies, N.; Fraser, J. L.; Galvin, F. C.; Hannah, D. R.; Haughan, A. F.; Jenkins, K.; Mack, S. R.; Pitt, W. R.; Ratcliffe, A. J.; Richard, M. D.; Sabin, V.; Sharpe, A.; Williams, S. C. Novel Indole Inhibitors of IMPDH from Fragments: Synthesis and Initial Structure–activity Relationships. *David W Robertson Meml. Issue* **2006**, *16*, 2539–2542.

(44) Pickett, S. D.; Sherborne, B. S.; Wilkinson, T.; Bennett, J.; Borkakoti, N.; Broadhurst, M.; Hurst, D.; Kilford, I.; McKinnell, M.; Jones, P. S. Discovery of Novel Low Molecular Weight Inhibitors of IMPDH via Virtual Needle Screening. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 1691–1694.

(45) Ziarek, J. J.; Liu, Y.; Smith, E.; Zhang, G.; Peterson, F. C.; Chen, J.; Yu, Y.; Chen, Y.; Volkman, B. F.; Li, R. Fragment-Based Optimization of Small Molecule CXCL12 Inhibitors for Antagonizing the CXCL12/CXCR4 Interaction. *Curr. Top. Med. Chem. 12*, 2727–2740.

(46) Lau, W.; Withka, J.; Hepworth, D.; Magee, T.; Du, Y.; Bakken, G.; Miller, M.; Hendsch, Z.; Thanabal, V.; Kolodziej, S.; Xing, L.; Hu, Q.; Narasimhan, L.; Love, R.; Charlton, M.; Hughes, S.; Hoorn, W.; Mills, J. Design of a Multi-Purpose Fragment Screening Library Using Molecular Complexity and Orthogonal Diversity Metrics. *J. Comput. Aided Mol. Des.* **2011**, *25*, 621–636.

(47) Carr, R. A. E.; Congreve, M.; Murray, C. W.; Rees, D. C. Fragment-Based Lead Discovery: Leads by Design. *Drug Discov. Today* **2005**, *10*, 987–992.

(48) Chen, I.-J.; Hubbard, R. Lessons for Fragment Library Design: Analysis of Output from Multiple Screening Campaigns. *J. Comput. Aided Mol. Des.* **2009**, *23*, 603–620.

(49) Barnard, J. M.; Downs, G. M. Chemical Fragment Generation and Clustering Software§. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 141–142.

(50) Miller, M. A. Chemical Database Techniques in Drug Discovery. *Nat Rev Drug Discov* **2002**, *1*, 220–227.

(51) Walters, W. P.; Stahl, M. T.; Murcko, M. A. Virtual Screening—an Overview. *Drug Discov. Today* **1998**, *3*, 160–178.

(52) Rognan, D. Development and Virtual Screening of Target Libraries. *World Synap. Mol. Basis Pathol. Drug Discov.* **2006**, *99*, 232–244.

(53) Tan, L.; Lounkine, E.; Bajorath, J. Similarity Searching Using Fingerprints of Molecular Fragments Involved in Protein−Ligand Interactions. *J. Chem. Inf. Model.* **2008**, *48*, 2308–2312.

(54) Grasberger, B. L.; Lu, T.; Schubert, C.; Parks, D. J.; Carver, T. E.; Koblish, H. K.; Cummings, M. D.; LaFrance, L. V.; Milkiewicz, K. L.; Calvo, R. R.; Maguire, D.; Lattanze, J.; Franks, C. F.; Zhao, S.; Ramachandren, K.; Bylebyl, G. R.; Zhang, M.; Manthey, C. L.; Petrella, E. C.; Pantoliano, M. W.; Deckman, I. C.; Spurlino, J. C.; Maroney, A. C.; Tomczuk, B. E.; Molloy, C. J.; Bone, R. F. Discovery and Cocrystal

Structure of Benzodiazepinedione HDM2 Antagonists That Activate p53 in Cells. *J. Med. Chem.* **2005**, *48*, 909–912.

(55) Liebeschuetz, J. W.; Jones, S. D.; Morgan, P. J.; Murray, C. W.; Rimmer, A. D.; Roscoe, J. M. E.; Waszkowycz, B.; Welsh, P. M.; Wylie, W. A.; Young, S. C.; Martin, H.; Mahler, J.; Brady, L.; Wilkinson, K. PRO_SELECT: Combining Structure-Based Drug Design and Array-Based Chemistry for Rapid Lead Discovery. 2. The Development of a Series of Highly Potent and Selective Factor Xa Inhibitors. *J. Med. Chem.* **2002**, *45*, 1221–1232.

(56) Zartler, E. R.; Sharpiro, M. J. *Fragment Based Drug Discovery - A Practical Approach*.

(57) Ciulli, A.; Abell, C. Fragment-Based Approaches to Enzyme Inhibition. *Chem. Biotechnol. Pharm. Biotechnol.* **2007**, *18*, 489–496.

(58) Hung, A. W.; Silvestre, H. L.; Wen, S.; Ciulli, A.; Blundell, T. L.; Abell, C. Application of Fragment Growing and Fragment Linking to the Discovery of Inhibitors of Mycobacterium Tuberculosis Pantothenate Synthetase. *Angew. Chem. Int. Ed.* **2009**, *48*, 8452–8456.

(59) Amaning, K.; Lowinski, M.; Vallee, F.; Steier, V.; Marcireau, C.; Ugolini, A.; Delorme, C.; Foucalt, F.; McCort, G.; Derimay, N.; Andouche, C.; Vougier, S.; Llopart, S.; Halland, N.; Rak, A. The Use of Virtual Screening and Differential Scanning Fluorimetry for the Rapid Identification of Fragments Active against MEK1. *Bioorg. Med. Chem. Lett.* **2013**, *23*, 3620–3626.

(60) Koshland, D. E. Application of a Theory of Enzyme Specificity to Protein Synthesis. *Proc. Natl. Acad. Sci.* **1958**, *44*, 98–104.

(61) Lo, M.-C.; Aulabaugh, A.; Jin, G.; Cowling, R.; Bard, J.; Malamas, M.; Ellestad, G. Evaluation of Fluorescence-Based Thermal Shift Assays for Hit Identification in Drug Discovery. *Anal. Biochem.* **2004**, *332*, 153–159.

(62) Dai, R.; Wilson, D. J.; Geders, T. W.; Aldrich, C. C.; Finzel, B. C. Inhibition of Mycobacterium Tuberculosis Transaminase BioA by Aryl Hydrazines and Hydrazides. *ChemBioChem* **2014**, *15*, 575–586.

(63) Lepre, C. A.; Moore, J. M.; Peng, J. W. Theory and Applications of NMR-Based Screening in Pharmaceutical Research. *Chem. Rev.* **2004**, *104*, 3641–3676.

(64) Venkitakrishnan, R.; Benard, O.; Max, M.; Markley, J.; Assadi-Porter, F. Use of NMR Saturation Transfer Difference Spectroscopy to Study Ligand Binding to Membrane Proteins. In *Membrane Protein Structure and Dynamics*; Vaidehi, N.; Klein-Seetharaman, J., Eds.; Methods in Molecular Biology; Humana Press, 2012; Vol. 914, pp. 47–63.

(65) Navratilova, I.; Hopkins, A. L. Fragment Screening by Surface Plasmon Resonance. *ACS Med. Chem. Lett.* **2010**, *1*, 44–48.

(66) Neumann, T.; Sekul, R. SPR Screening of Chemical Microarrays for Fragment-Based Discovery. In *Label-Free Technologies for Drug Discovery*; John Wiley & Sons, Ltd, 2011; pp. 41–55.

(67) Perspicace, S.; Banner, D.; Benz, J.; Müller, F.; Schlatter, D.; Huber, W. Fragment-Based Screening Using Surface Plasmon Resonance Technology. *J. Biomol. Screen.* **2009**, *14*, 337–349.

(68) Pellecchia, M.; Meininger, D.; Dong, Q.; Chang, E.; Jack, R.; Sem, D. NMR-Based Structural Characterization of Large Protein-Ligand Interactions. *J. Biomol. NMR* **2002**, *22*, 165–173.

(69) Bax, A.; Grzesiek, S. Methodological Advances in Protein NMR. *Acc. Chem. Res.* **1993**, *26*, 131–138.

(70) Aramini, J. M.; Rossi, P.; Huang, Y. J.; Zhao, L.; Jiang, M.; Maglaqui, M.; Xiao, R.; Locke, J.; Nair, R.; Rost, B.; Acton, T. B.; Inouye, M.; Montelione, G. T. Solution NMR Structure of the NlpC/P60 Domain of Lipoprotein Spr from Escherichia Coli: Structural Evidence for a Novel Cysteine Peptidase Catalytic Triad†. *Biochemistry (Mosc.)* **2008**, *47*, 9715–9717.

(71) Cole, S. T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S. V.; Eiglmeier, K.; Gas, S.; Barry, C. E.; Tekaia, F.; Badcock, K.; Basham, D.; Brown, D.; Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Hornsby, T.; Jagels, K.; Krogh, A.; McLean, J.; Moule, S.; Murphy, L.; Oliver, K.; Osborne, J.; Quail, M. A.; Rajandream, M.-A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, J.; Squares, R.; Squares, S.; Sulston, J. E.; Taylor, K.; Whitehead, S.; Barrell, B. G. Deciphering the Biology of Mycobacterium Tuberculosis from the Complete Genome Sequence. *Nature* **1998**, *393*, 537–544.

(72) Geitmann, M.; Elinder, M.; Seeger, C.; Brandt, P.; de Esch, I. J. P.; Danielson, U. H. Identification of a Novel Scaffold for Allosteric Inhibition of Wild Type and Drug Resistant HIV-1 Reverse Transcriptase by Fragment Library Screening. *J. Med. Chem.* **2011**, *54*, 699–708.

(73) Wilson, D. J.; Shi, C.; Duckworth, B. P.; Muretta, J. M.; Manjunatha, U.; Sham, Y. Y.; Thomas, D. D.; Aldrich, C. C. A Continuous Fluorescence Displacement Assay for BioA: An Enzyme Involved in Biotin Biosynthesis. *Anal. Biochem.* **2011**, *416*, 27–38.

(74) Geders, T. W.; Gustafson, K.; Finzel, B. C. Use of Differential Scanning Fluorimetry to Optimize the Purification and Crystallization of PLP-Dependent Enzymes. *Acta Crystallogr. Sect. F* **2012**, *68*, 596–600.

(75) Eliot, A. C.; Sandmark, J.; Schneider, G.; Kirsch, J. F. The Dual-Specific Active Site of 7,8-Diaminopelargonic Acid Synthase and the Effect of the R391A Mutation†. *Biochemistry (Mosc.)* **2002**, *41*, 12582–12589.

(76) Kabsch, W. XDS. *Acta Crystallogr. Sect. D* **2010**, *66*, 125–132.

(77) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. Phaser Crystallographic Software. *J. Appl. Crystallogr.* **2007**, *40*, 658–674.

(78) Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.;

Wilson, K. S. Overview of the CCP4 Suite and Current Developments. *Acta Crystallogr. Sect. D* **2011**, *67*, 235–242.

(79) Vagin, A. A.; Steiner, R. A.; Lebedev, A. A.; Potterton, L.; McNicholas, S.; Long, F.; Murshudov, G. N. REFMAC5 Dictionary: Organization of Prior Chemical Knowledge and Guidelines for Its Use. *Acta Crystallogr. Sect. D* **2004**, *60*, 2184–2195.

(80) Emsley, P.; Cowtan, K. Coot: Model-Building Tools for Molecular Graphics. *Acta Crystallogr. Sect. D* **2004**, *60*, 2126–2132.

(81) Finzel, B. C.; Akavaram, R.; Ragipindi, A.; Van Voorst, J. R.; Cahn, M.; Davis, M. E.; Pokross, M. E.; Sheriff, S.; Baldwin, E. T. Conserved Core Substructures in the Overlay of Protein–Ligand Complexes. *J. Chem. Inf. Model.* **2011**, *51*, 1931–1941.

(82) Kranz, J. K.; Schalk-Hihi, C. Chapter Eleven - Protein Thermal Shifts to Identify Low Molecular Weight Fragments. In *Methods in Enzymology*; Lawrence C. Kuo, Ed.; Academic Press, 2011; Vol. Volume 493, pp. 277–298.

(83) Petros, A. M.; Dinges, J.; Augeri, D. J.; Baumeister, S. A.; Betebenner, D. A.; Bures, M. G.; Elmore, S. W.; Hajduk, P. J.; Joseph, M. K.; Landis, S. K.; Nettesheim, D. G.; Rosenberg, S. H.; Shen, W.; Thomas, S.; Wang, X.; Zanze, I.; Zhang, H.; Fesik, S. W. Discovery of a Potent Inhibitor of the Antiapoptotic Protein Bcl-xL from NMR and Parallel Synthesis. *J. Med. Chem.* **2005**, *49*, 656–663.

(84) Hartshorn, M. J.; Murray, C. W.; Cleasby, A.; Frederickson, M.; Tickle, I. J.; Jhoti, H. Fragment-Based Lead Discovery Using X-Ray Crystallography. *J. Med. Chem.* **2004**, *48*, 403–413.

(85) Hennig, M.; Ruf, A.; Huber, W. Combining Biophysical Screening and X-Ray Crystallography for Fragment-Based Drug Discovery. In *Fragment-Based Drug Discovery and X-Ray Crystallography*; Davies, T. G.; Hyvönen, M., Eds.; Topics in Current Chemistry; Springer Berlin Heidelberg, 2012; Vol. 317, pp. 115–143.

(86) Silvestre, H. L.; Blundell, T. L.; Abell, C.; Ciulli, A. Integrated Biophysical Approach to Fragment Screening and Validation for Fragment-Based Lead Discovery. *Proc. Natl. Acad. Sci.* **2013**, *110*, 12984–12989.

(87) Angulo, J.; Enríquez-Navas, P. M.; Nieto, P. M. Ligand–Receptor Binding Affinities from Saturation Transfer Difference (STD) NMR Spectroscopy: The Binding Isotherm of STD Initial Growth Rates. *Chem. – Eur. J.* **2010**, *16*, 7803–7812.

(88) Evans, P. Scaling and Assessment of Data Quality. *Acta Crystallogr. Sect. D* **2006**, *62*, 72–82.

(89) Tan, Y. S.; Śledź, P.; Lang, S.; Stubbs, C. J.; Spring, D. R.; Abell, C.; Best, R. B. Using Ligand-Mapping Simulations to Design a Ligand Selectively Targeting a Cryptic Surface Pocket of Polo-Like Kinase 1. *Angew. Chem. Int. Ed.* **2012**, *51*, 10078–10081.

(90) Murray, C. W.; Carr, M. G.; Callaghan, O.; Chessari, G.; Congreve, M.; Cowan, S.; Coyle, J. E.; Downham, R.; Figueroa, E.; Frederickson, M.; Graham, B.; McMenamin, R.; O'Brien, M. A.; Patel, S.; Phillips, T. R.; Williams, G.; Woodhead, A. J.; Woolford, A. J.-A. Fragment-Based Drug Discovery Applied to Hsp90.

Discovery of Two Lead Series with High Ligand Efficiency. *J. Med. Chem.* **2010**, *53*, 5942–5955.

(91) Ferenczy, G. G.; Keserű, G. M. Thermodynamics of Fragment Binding. *J. Chem. Inf. Model.* **2012**, *52*, 1039–1045.

(92) Surade, S.; Ty, N.; Hengrung, N.; Lechartier, B.; Cole, S. T.; Abell, C.; Blundell, T. L. A Structure-Guided Fragment-Based Approach for the Discovery of Allosteric Inhibitors Targeting the Lipophilic Binding Site of Transcription Factor EthR. *Biochem. J.* **2014**, *458*, 387–394.

(93) Cheng, L.; Pettersen, D.; Ohlsson, B.; Schell, P.; Karle, M.; Evertsson, E.; Pahlén, S.; Jonforsen, M.; Plowright, A. T.; Boström, J.; Fex, T.; Thelin, A.; Hilgendorf, C.; Xue, Y.; Wahlund, G.; Lindberg, W.; Larsson, L.-O.; Gustafsson, D. Discovery of the Fibrinolysis Inhibitor AZD6564, Acting via Interference of a Protein–Protein Interaction. *ACS Med. Chem. Lett.* **2014**, *5*, 538–543.

(94) Seth, P. P.; Miyaji, A.; Jefferson, E. A.; Sannes-Lowery, K. A.; Osgood, S. A.; Propp, S. S.; Ranken, R.; Massire, C.; Sampath, R.; Ecker, D. J.; Swayze, E. E.; Griffey, R. H. SAR by MS: Discovery of a New Class of RNA-Binding Small Molecules for the Hepatitis C Virus: Internal Ribosome Entry Site IIA Subdomain. *J. Med. Chem.* **2005**, *48*, 7099–7102.

(95) Fan, C.; Vederas, J. C. Synthesis and Structure-Activity Relationships of O-Sulfonamido-Arylhydrazides as Inhibitors of Ll-Diaminopimelate Aminotransferase (ll-DAP-AT). *Org. Biomol. Chem.* **2012**, *10*, 5815–5819.

(96) Hopkins, A. L.; Keseru, G. M.; Leeson, P. D.; Rees, D. C.; Reynolds, C. H. The Role of Ligand Efficiency Metrics in Drug Discovery. *Nat Rev Drug Discov* **2014**, *13*, 105–121.

(97) Schuttelkopf, A. W.; van Aalten, D. M. F. PRODRG: A Tool for High-Throughput Crystallography of Protein-Ligand Complexes. *Acta Crystallogr. Sect. D* **2004**, *60*, 1355–1363.

(98) Kaneko, M.; Kontani, Y.; Kikugawa, M.; Tamaki, N. Inhibition of D-3-Aminoisobutyrate-Pyruvate Aminotransferase by 5-Fluorouracil and A-Fluoro-B-Alanine. *Biochim. Biophys. Acta BBA - Protein Struct. Mol. Enzymol.* **1992**, *1122*, 45–49.

(99) Hayashi, H.; Wada, H.; Yoshimura, T.; Esaki, N.; Soda, K. Recent Topics in Pyridoxal 5'-Phosphate Enzyme Studies. *Annu. Rev. Biochem.* **1990**, *59*, 87–110.

(100) Barends, T. R. M.; Domratcheva, T.; Kulik, V.; Blumenstein, L.; Niks, D.; Dunn, M. F.; Schlichting, I. Structure and Mechanistic Implications of a Tryptophan Synthase Quinonoid Intermediate. *ChemBioChem* **2008**, *9*, 1024–1028.

(101) Lightcap, E. S.; Hopkins, M. H.; Olson, G. T.; Silverman, R. B. Time-Dependent Inhibition of Γ-Aminobutyric Acid Aminotransferase, by 3-Hydroxybenzylhydrazine. *Bioorg. Med. Chem.* **1995**, *3*, 579–585.

(102) Ejim, L. J.; Blanchard, J. E.; Koteva, K. P.; Sumerfield, R.; Elowe, N. H.; Chechetto, J. D.; Brown, E. D.; Junop, M. S.; Wright, G. D. Inhibitors of Bacterial Cystathionine B-Lyase: Leads for New Antimicrobial Agents and Probes of Enzyme Structure and Function. *J. Med. Chem.* **2007**, *50*, 755–764.

(103) Zlitni, S.; Ferruccio, L. F.; Brown, E. D. Metabolic Suppression Identifies New Antibacterial Inhibitors under Nutrient Limitation. *Nat Chem Biol* **2013**, *advance online publication*.

(104) Niesen, F. H.; Berglund, H.; Vedadi, M. The Use of Differential Scanning Fluorimetry to Detect Ligand Interactions That Promote Protein Stability. *Nat Protoc.* **2007**, *2*, 2212–2221.

(105) Ji, H.; Silverman, R. B. Case Study 3: Fragment Hopping to Design Highly Potent and Selective Neuronal Nitric Oxide Synthase Inhibitors. In *Scaffold Hopping in Medicinal Chemistry*; Wiley-VCH Verlag GmbH & Co. KGaA, 2013; pp. 279–296.

(106) Böhm, H.-J.; Flohr, A.; Stahl, M. Scaffold Hopping. *Drug Discov. Today Technol.* **2004**, *1*, 217–224.

(107) Sun, H.; Tawa, G.; Wallqvist, A. Classification of Scaffold-Hopping Approaches. *Drug Discov. Today* **2012**, *17*, 310–324.

(108) Jenkins, J. L.; Glick, M.; Davies, J. W. A 3D Similarity Method for Scaffold Hopping from Known Drugs or Natural Ligands to New Chemotypes. *J. Med. Chem.* **2004**, *47*, 6144–6159.

(109) Ostermann, N.; Ruedisser, S.; Ehrhardt, C.; Breitenstein, W.; Marzinzik, A.; Jacoby, E.; Vangrevelinghe, E.; Ottl, J.; Klumpp, M.; Hartwieg, J. C. D.; Cumin, F.; Hassiepen, U.; Trappe, J.; Sedrani, R.; Geisse, S.; Gerhartz, B.; Richert, P.; Francotte, E.; Wagner, T.; Krömer, M.; Kosaka, T.; Webb, R. L.; Rigel, D. F.; Maibaum, J.; Baeschlin, D. K. A Novel Class of Oral Direct Renin Inhibitors: Highly

Potent 3,5-Disubstituted Piperidines Bearing a Tricyclic P3–P1 Pharmacophore. *J. Med. Chem.* **2013**, *56*, 2196–2206.

(110) Jenkins, J. L. Feature Point Pharmacophores (FEPOPS). In *Scaffold Hopping in Medicinal Chemistry*; Wiley-VCH Verlag GmbH & Co. KGaA, 2013; pp. 155–174.

(111) Pflugrath, J. The Finer Things in X-Ray Diffraction Data Collection. *Acta Crystallogr. Sect. D* **1999**, *55*, 1718–1725.

(112) Adams, P. D.; Afonine, P. V.; Bunkoczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L.-W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H. PHENIX: A Comprehensive Python-Based System for Macromolecular Structure Solution. *Acta Crystallogr. Sect. D* **2010**, *66*, 213–221.

(113) Lebedev, A. A.; Young, P.; Isupov, M. N.; Moroz, O. V.; Vagin, A. A.; Murshudov, G. N. JLigand: A Graphical Tool for the CCP4 Template-Restraint Library. *Acta Crystallogr. Sect. D* **2012**, *68*, 431–440.

(114) Sensi, P.; Margalith, P.; Timbal, M. Rifomycin, a New Antibiotic; Preliminary Report. *Farm. Sci.* **1959**, *14*, 146–147.